The Psychological Record

a quarterly journal in theoretical

and experimental psychology

■ ■ ■ ■ ■ ■ ■ ■

CONTENTS

■ ■ ■ ■ ■ ■ ■ ■

THE PSYCHOLOGICAL RECORD is a non-profit publication. It is published quarterly in January, April, July, and October, at Denison University, Granville, Ohio. Subscription price is $4.00 a year (APA members—$3.00; students—$1.50).

With the permission of the Principia Press, Inc., THE PSYCHOLOGICAL RECORD is a continuation of the journal formerly published under this title. Publication of THE PSYCHOLOGICAL RECORD was resumed in January, 1956.

As presently organized THE PSYCHOLOGICAL RECORD publishes both theoretical and experimental articles, commentary on current developments in psychology, and descriptions of research planned or in progress. The journal is designed to serve a *critical function in psychology*. It therefore favors the publication of papers that develop new approaches to the study of behavior and new methodologies, and which undertake critiques of existing approaches and methods.

Articles should be prepared according to the form suggested for APA publications (*APA Publication Manual*) and submitted in duplicate to the Editor. The author cost per page is $3.00. There is an additional author charge for cuts and special composition. Reprints are available at cost.

A BRIEF HISTORY OF EDUCATIONAL PSYCHOLOGY

ROBERT I. WATSON[1]
Northwestern University

A succinct history of educational psychology is not available. At present we must depend upon material scattered in various sources. A few illustrations of the contents of these sources might be mentioned. Boring (1950), in his history of experimental psychology, devotes a few pages to educational psychology as an aspect of functional psychology. Murphy (1949), in his introduction to the history of modern psychology, makes incidental reference to educational psychology in his discussions of the acquisition of skill and of child psychology. Short histories of other fields such as child psychology (Dennis, 1949), and clinical psychology (Watson, 1953), are useful, but do not directly apply to the issue at hand. The account of the impact of psychology upon education by Burt (1957), although admirable, is written for a British audience and published in a not too accessible source. A few textbooks in educational psychology devote a page or two to this topic. The fullest treatments of educational psychology probably are to be found in various histories of education. These accounts are generally unsatisfactory for reasons unnecessary to go into.

A summary of this history of educational psychology would seem to be desirable. The account of the history of educational psychology is organized in the following fashion. The period before the emergence of psychology as a separate discipline is reviewed. The beginnings of educational psychology during the years from 1880 to 1900 are then considered with particular emphasis upon the contributions of the founders of the field. The period of the incubation of educational psychology (1900-1918) is next presented both in terms of the leaders and of certain major problems which received emphasis. The recent past, defined for the purposes of this paper as extending from 1918 to 1941, then receives consideration. Some reflections on the significance of these findings close the paper.

BEFORE EDUCATIONAL PSYCHOLOGY

Before the emergence of psychology as a discipline in its own right, and of educational psychology as a branch of that discipline, speculation and observation concerning the relation of human nature to the educational process were not uncommon. This period began in classical Greece. If one examines the fragments of the thinking of the pre-

Socratic philosophers that have been preserved there is to be found a variety of comments relevant to this relation. Democritus (Freeman, 1953), who was in his prime early in the fifth century B.C., will serve as an example. He wrote not only about the advantages of education but also considered the influence of the home upon the child. He held that self control of the father served to teach the children. Stress was laid upon such matters as training in the management of property by sharing it with the children. He also commented that some qualities present in the parents result in their opposite appearing in the child, giving as an instance the parental practice of excessive thrift leading to extravagance in the children.

Of Plato and of Aristotle, who flourished in the fourth century B.C., we have much fuller accounts of their views of education and its relation to psychological factors. The scope of their thinking might seem to some persons today to be surprisingly complex and broad. Scattered throughout their works Aristotle and Plato discuss the following: the ends of education; the ideal of the educated man, the disadvantages of being educated; the kinds of education that are appropriate to different kinds of people; the training of the body, and the cultivation of bodily skills; the formation of a good character; the possibilities and limits of moral education; the influence of the family in this training; the role of the state in moral education; the effect upon character of music, poetry and the other arts; the profession of teaching, and the relation of teacher and student; the means and methods of teaching; the nature of learning; the order of learning; the emotional aspects of learning; learning apart from teachers, and the acquisition of techniques (Adler, 1952).

Aristotle's psychological views relevant to educational matters are presented more systematically, less fancifully, and in greater detail than are those of Plato. Two aspects of Aristotle's thinking especially worthy of note are that his writings served to encourage a faculty psychology and that he emphasized the intellectual processes. Aristotle laid the foundation for the main psychological doctrines taught and accepted by the learned world for the next 2000 years. In the earlier darker years such knowledge virtually disappeared, only to be revived again.

Throughout the centuries, the formulations of Aristotle have been modified in particulars. The most noteworthy of these modifications took place when the medieval scholastic thinkers attempted to reconcile Christian doctrine with Aristotelian teaching. These efforts reached their climax in the work of Thomas Aquinas (1911-1920, 1953) in the thirteenth century.

Descartes in the seventeenth century introduced powerful, reasoned support for the doctrine that ideas which are innate, rather than those arising from experience, are the basis of true knowledge (Brett, 1921).

It was John Locke (Brett, 1921), a seventeenth century empiricist, who initiated the first wide and continued protest against the form of faculty psychology current in his time. He did not abandon faculties altogether, but argued that the faculties were not real things in the soul that performed the actions indicated by their various names. Locke's *Essay concerning Human Understanding* (first published in 1690 and including in the fourth edition, published in 1700, a chapter entitled "Of the Association of Ideas") is an attempt to combat the Cartesian theory of innate ideas. Locke insisted that, at birth, the human mind was not preformed and ready to function, but potentially sensitive to impressions from the external world through the senses. Through the empirical process of experience he receives sensory experiences from which simple experiences are built into complex ideas through the internal activity of reflection. Ideas, then, were not already present but arose from sensory impressions and from the process of reflection. Thus arise all knowledge, all values, and all morality. This learning through experience came to be known as empiricism. Other psychologists that came after him took the position that the source of ideas was sensory experiences alone and rejected his faculty of reflection. From his work and from the work of those that followed after a thorough-going non-faculty point of view began to emerge—association psychology (Warren, 1921). Ironically enough, this associationist point of view also goes back to Aristotle who discussed the lawfulness of association and spoke of the importance of similarity, contrast, and contiguity. The struggle over faculty psychology and over association psychology, eventually sharply separated, continued for about two hundred years after Locke's publications at the turn of the eighteenth century.

In the course of the development of the doctrine of faculties its ramifications, subtleties and nuances were as detailed, contradictory and as hotly contested as any of the ideas of Western man. The doctrine may be simplified to the statement that most commonly the mind was considered as three interdependent (or, conversely, dependent), sets of powers or capacities, the first and foremost of which was understanding, reasoning or intellect, another was feelings, desires, emotions and appetites, and the third was the will or volition.

At the beginning of the nineteenth century, although continuing the faculty psychology, Pestalozzi changed the sphere of teacher training by placing an emphasis upon education as a process of drawing out of the individual. This, in turn, required a method to be sought to minimize that of the memorizing and the hearing of recitations which was then prevalent. Proceeding from the concrete to the abstract, attending to the harmoniousness of development, and searching for the laws of development characterized his point of view (Cubberly, 1920).

Faculty psychology held an unquestioned firm grip upon education in the United States from the 1830's to almost the end of the last century (Butts & Cremin, 1953). Faculty psychology set the limits and form

of education. It provided the basic justification for the method of mental discipline, giving first place to a liberal education, in the original sense of the term, and served as an obstacle to the appearance of new and potentially useful subjects in the curriculum. The faculties of the students were to be exercised by use of those subjects of instruction, which, in the words of Day, President of Yale in the early decades of the century, are "best calculated to teach the art of fixing the attention, directing the train of thought, analyzing a subject proposed for investigation, following, with accurate discrimination, the course of argument; balancing nicely the evidence presented to the judgment; awakening, elevating, and controlling the imagination; arranging, with skill, the treasures which memory gathers; rousing and guiding the powers of genius . . ." (1829, Pp. 300-301). The concentration on the intellectual faculties to the virtual exclusion of the faculties of feeling and will is obvious.

The next great advance in educational psychology came about after the middle part of the century when the neglected work of Johann Friedrich Herbart became influential. Herbart was a German professor, who, according to Eby and Arrowood (1952), was the first to formulate an approach to education based directly and avowedly upon psychology. However, as his predecessors before him, he was primarily philosophical in aim, content, and method. He specifically denied the possibility of experimenting upon the mind. Nevertheless, he rejected the faculty psychology of his predecessors and stressed the importance of interest and apperception and apperceptive mass. Herbart (Asmus, 1957), not only opposed the faculty psychologists of his time, but also considered the human personality as a dynamic and individually structured system of forces. Ideas are active and struggle with one another for a place in consciousness. His work in education was carried on adjunct to his major task of philosophy, and at the time of his death in 1841 and for some years thereafter went practically unnoticed (Eby and Arrowood, 1952). In 1868 a professor at Leipzig, Ziller, published a book calling attention to his work. This book was widely read. Herbart's influence spread rapidly thereafter with his ideas elaborated and advanced as the primary basis for school instruction. In the 1890's the influence of Herbart upon educational psychology in the United States was all-embracing in nature.

In England, meanwhile, two important books were published. The volume of Alexander Bain, *Education as a Science,* was published in 1878 and emphasized an associationist view. James Sully's *Outlines of Psychology with Special Reference to the Theory of Education,* was published in 1884 and emphasized a faculty point of view. Because of the all-embracing nature of these conflicting two views at that time, it is quite fitting that mention of these two contrasting books should close the account of the period preceding the advent of educational psychology as an aspect of psychology.

THE BEGINNINGS OF EDUCATIONAL PSYCHOLOGY: 1880-1900

There is considerable justification for placing the beginnings of educational psychology in the early years of the 1880's. During this time Galton published the first experimental investigation of associationism (1879), and began giving to the public a series of physical measurements and tests of reaction time and sensory acuity (1884). Hall, meanwhile, (1883), published the first of his papers using the questionnaire to investigate the minds of children. In 1885 Ebbinghaus published his study of memory, *Ueber das Gedächtnis* (1913). In the span of six years events important in the founding of objective measurement, child psychology and learning all took place. In the present perspective they are also seen to be crucial aspects of educational psychology.

Francis Galton. Sir Francis Galton, the oldest of the founders of educational psychology, was influenced by what Warren (1921) called the evolutionary associationist tradition, particularly the views espoused by Lewes and Spencer. Associationism under the influence of Darwin had taken on a more developmental cast. As was just mentioned, it was Galton who published in 1879 the first experimental investigations of associationism (Warren, 1921). The three studies he reported in this paper were concerned with the free play of association, the reaction time necessary to produce associations, and perseveration, or the tendency to repeat the same associations. His work on association was repeated in Wundt's laboratory the following year. Seven years later, also at the Leipzig laboratory, James McKeen Cattell (Boring, 1950) performed his well known series of studies on reaction time in relation to various mental processes. However, these contributions from Wundt's pupils had only an indirect connection with further developments in associationism.

To return to Galton, his work was guided by two basic Darwinian assumptions—that there is a resemblance in innate constitution of offspring to their parents, and that the innate constitution of the offspring would, because of spontaneous variation, bring about to some degree differences both from their parents and from one another (Burt, 1957). In 1869 he published his study, *Hereditary Genius*, which showed that eminent men in England tend to be related to one another. He conducted many other studies on mental inheritance among persons of ability and studied the resemblance of twins. Galton became firmly convinced of the importance of hereditary differences. With this emphasis upon hereditary differences, a wide-spread realization of the importance of individual differences, in the psychological sense, came into being. This interest in the nature-nurture problem took the form not only of individual differences conceived as being innate, but also seen as being due to environmental differences. Galton's enthusiasm for problems in this area led to his making an important advance in the history of statistics—the idea of correlation. This was followed up by

Karl Pearson who developed the measure of relationships that bears his name. The concept was that of Galton; the technique was that of Pearson (Walker, 1958). Further contributions from Galton took the form of the development and use of psychological tests, rating scales and questionnaires (Burt, 1957). His most important theoretical contribution was the distinction in the "structure of mind" between a general broad ability, or intelligence, and special abilities entering only into narrower ranges of activity. His conceptual distinction, his statistical tool and his testing devices bore rich fruit in studies of factor analysis.

G. *Stanley Hall.* Although the influence of G. Stanley Hall has been most directly felt in child psychology, he deserves mention in a discussion of the history of educational psychology. From the 1880's onward as a well known psychologist, as an academic leader in his capacity of President of Clark University, and as a promoter of worthy causes in psychology he exerted an influence which overrode considerably the theoretical weaknesses inherent in his own research. His conceptual scheme was based upon theories of biological evolution and recapitulation, in which the mind was seen as growing through a series of stages, more or less corresponding to those through which it was alleged early man had gone. Although enjoying considerable vogue at the time, today there remains very little trace of this point of view. A modern critical evaluation of his work may be found in Eby and Arrowood (1952).

He became interested in the possibilities of the questionnaire which he had found being used earlier in Berlin during the 1870's. With his characteristic flair for creating enthusiasm in others he inspired much research designed to discover the contents of children's minds. In the next decade his first paper appeared that used the questionnaire for the purpose (1883). This was followed by a whole series of studies, some carried out by him and his students but more carried out by others. The use of the technique became a fad, being applied by parents, teachers and others and a rash of uncritical, poorly designed, superficial, statistically unsophisticated papers appeared. This use of the questionnaire was the major impetus for the development of the so-called child study movement, with societies formed here and abroad (Bradbury, 1937). Hall, himself, lost interest in the movement before it had run its course, but through his inspiration and through his providing, beginning in 1891, a vehicle for publication, in the *Pedagogical Seminary* (now the *Journal of Genetic Psychology*) he was the major impetus for the movement. The period of most intense activity of this movement was from about 1890 to 1915. Eventually, the child study movement, in its original sense, fell of its own weight, and yet, as Bradbury (1937) indicates, the movement made its positive contributions in the increased recognition of the importance of the empirical study of the child, an increased critical evaluation of research, and a

recognition of the importance of the study of childhood development itself.

William James. William James, to use Boring's expression (1948), was a "self starter." Influenced to some degree by many individuals he had no mentor but himself. His book, *Principles of Psychology,* published in 1890, had a profound effect upon the development of psychology as a natural science. Drawing upon evolutionary biology, physiology, associationism, and all of the other scattered intellectual current of his day, in this book he carved out the beginnings of the so-called functional approach to psychology. In it, he saw the mind in terms of use in survival and in competition. Critical of the crudities of materialistic physiological psychologists and of the brass instrument psychology, and with a vivid, even epigrammatic, style he did much to win from teachers and others an appreciation of what, to a considerable extent, had been heretofore generally regarded as a curiously lifeless subject—the mind (Burt, 1957). The book, itself, is free of specific reference to educational applications. Rather, it served as a source for others to draw implications for education. It performed what Morris (1950) so aptly called its particular function, that of a watershed, initiating the flow of the future. One of the lessons to be drawn from Allport's discussion (1943) of the paradoxical (and often contradictory) positions that James took on psychological issues is that his work is a gigantic projective reservoir of ideas. One goes to James for support on a matter of concern to psychology and finds it put forth enthusiastically and convincingly. If he does not look for contrary pronouncements, disregards them, disagrees with them, or harmonizes them, he can come away refreshed and supported. James also contributed to educational psychology through his ability as a public speaker and popularizer. Despite its mundane title, his *Talks to Teachers,* published in 1899, is a brilliant book, and is crammed with still pertinent practical suggestions on educational matters given in his usual vivid style.

James McKeen Cattell. The early work of James McKeen Cattell on individual differences in reaction time in Wundt's laboratory has already been mentioned. This research, recognized as not being in the tradition of the Leipzig Laboratory, was prophetic of Cattell's later attitude. He broke away from the Wundtian outlook in that he laid stress on individual differences, and vigorously defended the practical implications of psychological findings. His student days at Leipzig are so much a part of the folklore of psychology that it is sometimes overlooked that he was thereafter a lecturer at Cambridge University where he was much influenced by the man he said was "the greatest man whom I have known," Francis Galton (Roback, 1952). Cattell, on his return from Germany, served as Professor of Psychology at the University of Pennsylvania and then moved to Columbia University where he was

head of the Laboratory and of the Department for 26 of the years in which psychology was shaping itself.

Much of his influence on educational psychology was of a more general character than was that of some of the other pioneers. He worked assiduously toward promoting all applications of psychology, including educational psychology. He organized journals, (along with J. Mark Baldwin), (*The Psychological Review, Psychological Monographs,* and *Psychological Index*) served as an editor, (*Science, American Men of Science, Scientific Monthly,* and *School and Society*) and was a founder of the Psychological Corporation.

It was in the area of mental tests that he made his most direct contribution to educational psychology. While still at the University of Pennsylvania, he had administered mental tests to the students, and indeed, in describing the tests used in the article (1890), coined the term, "mental tests." He continued his testing program when he moved to Columbia. Data were collected for several years and a monograph on the results was prepared by Clark Wissler (1901). The correlations between the tests and academic class standing were hardly better than zero. Moreover, the tests did not correlate appreciably better among themselves. This was in sharp contrast to the substantial correlations among the academic subjects. The results, then, seemed to lead nowhere so far as the usefulness of the tests were concerned. An earlier study by Sharp (1899) had appeared from Titchener's laboratory at Cornell University with similar negative results. These two studies did much to make psychologists lose interest in the topic for the next few years. The tests that had been used were laboratory sensory and motor tasks—memory for ideas, association, memory span, reaction time, perceptual discrimination, and the like. They were not organized into scales but each test was treated separately. At the time it was not appreciated that these devices were not suitable as educational measures. It was not until Americans became familiar with the work of Alfred Binet that interest in mental testing revived.

Alfred Binet. The next major contributor to the founding of educational psychology is Alfred Binet, who developed the first widely used individual intelligence test. Binet introduced objectivity into an area which, before his work, had lent itself poorly to direct vigorous investigation. In 1890 he has become associated with the Sorbonne and its laboratory of physiological psychology, an association which continued until his death in 1911. At first he had worked on problems of abnormal psychology, but from about 1887 onwards his research was conducted primarily in the schools of Paris and its suburbs. In collaboration with V. Henri (1893), he had contended in the very first issue of *L'Année Psychologique,* in 1895, that individual psychology was much more significantly studied by investigation of the complex mental processes than by tests of sensory discrimination and reaction tests. Pursuing this aim in the decade that followed this paper, Binet carried on

many studies of the complex processes, such as tests of memory of words and of sentences, descriptions of objects and of pictures, and attention. The specific task that set him upon the development of the Binet Scale was the necessity of having an instrument whereby Parisian school children suspected of mental retardation could be identified so that they might be given special training. With the assistance of Theophile Simon he developed the first Binet Scale (1905).

When Binet and Simon were working on their first Scale there was sufficient work already accomplished to lead them to refer to the widespread use of tests. The fundamental concept they followed in its construction was that of what Binet called a metrical scale—tests with items arranged in order of increasing difficulty, with each test corresponding to a different mental level. Deliberately they selected a considerable variety of tests in order to tap various forms of mental capacity. Among the 30 tests in the 1905 scale there were measures of visual coordination, execution of simple orders, knowledge of objects, repetition of sentences, giving differences between pairs of familiar objects, repetition of digits, and making distinctions between abstract forms. They arranged the tests, on the basis of preliminary findings, in an ascending order of difficulty and administered the tentative scale to normal and subnormal children. Norms for children of ages three, five, seven, nine, and eleven were derived from this. The concept of mental age was clearly grasped but not worked out in practice until their 1908 and 1911 scales which further refined standardization of the instrument. (The introduction of the Binet-Simon Scale in the United States is considered later.)

John Dewey. The sheer diversity of the many contributions of John Dewey has, to some extent, obscured his very real gift to educational psychology (Thomas & Schneider, 1927). His psychological contributions, since they are imbedded in the matrix of his other interests, sometimes are not recognized for what they are. Grissman (1942) for example, in order to formulate a resume of the psychology of Dewey, had to search most of his most important works ostensibly concerned with other problems than those of psychology. The fact that Dewey performed little or nothing in the way of data research also strengthened the tendency to minimize his furtherance of educational psychology.

It was during the period 1894-1904, when head of the Department of Philosophy and Education of the University of Chicago, that he performed his more distinctively psychological contributions. Later, when he was Professor of Philosophy at Columbia University, his interests were exclusively in the philosophical and social spheres.

Although preceded by James, Dewey helped to found functionalism as a school of psychology in the late '90's. Along with Tufts, Mead, Moore, and Angell, there came into being the Chicago school of philosophical pragmatists and psychological functionalists. (Dewey, himself,

preferred the term "instrumentalism" over pragmatism.) Dewey criticized both the "idealistic" views of faculty psychology and the "realistic" views of a purely biological psychology for their neglect of the social facets of human nature. Both of these contrasting views saw the individual child in isolation. He insisted that learning and education were social in character. The child was conceived by him to be active in the process of growth, and not the passive recipient of environmental influences. In this connection it is important to note contrary to a widely held opinion that he stressed not only the conditions which make for adaptation of the environment to the individual child, but also the adapting of the child to that environment.

Since Dewey's psychological views are perhaps not as well known as those of some of the other pioneers a short description is appropriate. Dewey's fullest expression of functional psychology came in his article, "The Reflex Arc Concept in Psychology" (1896). He wanted it to be seen that, just as the stimulus calls forth the response, so too does the response bear on the stimulus, and that the whole reflex arc, instead of being a closed unit, is but a link in a chain of preceding and succeeding arcs. This chain gives us an effective instrument for effecting successful coordinations which aid the organism in its efforts toward the attainment of a goal. This is an adaptive process. The basic unit of Dewey's psychology was habit (Crissman, 1942). Habits are acquired dispositions which are dynamic and persistent. They form the basis of impulses, emotions, motives, desires, perceptions, imagination, thought, meaning, object, mind, consciousness, and self. Many of Dewey's contributions had been to the area between the philosophy of education and psychology in which he would draw upon both fields. In *How We Think* (1933) he discussed the use of intelligence and an intellectual method assimilated to the behavior of the individual as he solves his problems. The individual needs to acquire those habits through which his thinking comes under his control. His thought becomes an instrument of experience.

Dewey's most closely related educational contributions deserve mention. He was the intellectual father of progressive education, with his emphasis upon personal interests, social factors, and practical activities. The appearance of differentiated curricula, the activity program, elective subjects, learning readiness programs, and an increased concern with the individual student, all attest to the changing views of educational functions in which progressive education was greatly interested. Although Dewey was important in these developments, he was not alone, nor was he guilty of promulgating what later came to be regarded as the excesses of that movement. Progressive education had its roots in psychology. As Krugman (1948) reminds us, progressive education consists essentially of the application of mental hygiene to education.

It was through the work of these pioneers that educational psy-

chology was founded. Although working upon different problems and using diverse approaches, they had in common a desire for objectivity and a conviction that through quantitative research and measurement, educational psychology could be placed upon a firm footing. After these pioneers, and often as their students, came the first and subsequent generations of educational psychologists.

Closely Related Developments. Of necessity, this account of events related to the development of educational psychology occurring during these years must be very condensed. Attention will be confined to the development of, and the need for, courses in educational psychology; the beginning of the university study of education; the appearance of graduate schools; educational research through the measurement of achievement; and the beginnings of the psychology of learning.

Turning to courses in educational psychology, although it was not the first public normal school, the school at Oswego, N. Y., established in 1863, is important in that from the onset it had a course entitled "Child Study" (Crabb, 1926). In the years that followed the normal schools or departments of education being established routinely included courses in child study or educational psychology. The first of these terms was popularized by Hall in the nineties, but after publication of Thorndike's *Educational Psychology* (1903) the latter term was the one preferred.[2]

In educational circles in the late 1880's the need for courses in educational psychology was beginning to be voiced explicitly by educators. In 1888 in San Francisco the meeting of the National Educational Association devoted considerable attention to educational psychology and its role in the training of teachers. Blair (1948) tells us that Parr, for example, after speaking of the fact that, all too often, educational psychology is taken to mean merely the study of general psychology with stray observations about children, went on to indicate that eventually it will be the subject arising from applying the principles of general mental science to the conscious process of development. He indicates that at the same meeting, Baldwin stressed the value of educational psychology for placing the teacher in a position where he can better understand the pupil, and hence, more intelligently lead him. Similarly, Hodgin considered educational psychology imperatively necessary.

Turning to the university study of education, the first permanent chair in education, a Professorship of Philosophy and Education, was established at the State University of Iowa in 1873 (Cubberly, 1920). Up to 1890 less than a dozen chairs in education has been established in the United States. The work of these professors was largely limited to historical and philosophical studies. Thereafter, the number began

---

[2]The first publication to bear the title *Educational Psychology*, otherwise unimportant, was a booklet published in 1886 by Louisa Hopkins (Roback, 1952).

to increase rapidly and the nature of their research to change in the direction of the collection of quantitative data.

What was apparently the first department of education, the Department of the Science and Art of Teaching, was founded at the University of Michigan in 1879. This was followed by the establishment of a Department of Pedagogy at the University of Washington in 1881 and similar departments at the University of North Carolina and at the Johns Hopkins University in 1884. What was to become one of the foremost university training schools for teachers in the United States, Teachers College of Columbia University, was founded in 1888. Another leading school of education, that of the University of Chicago, was organized in 1900. In 1896-1897 the United States Commissioner of Education reported that 220 out of 432 institutions of higher learning offered courses in pedagogy.

Meanwhile, graduate schools in the modern sense were in the process of development. Before 1850 in the United States graduate education was a sporadic, unorganized affair (Little, 1960). The graduate schools of Harvard and Yale, already in existence, were very much overshadowed by their powerful undergraduate colleges to which they were subordinated (Eby & Arrowood, 1952). The first earned Ph.D. degree in the United States was conferred by Yale University in 1861. Johns Hopkins opened its doors as a graduate school in 1876. For the first time the ideal of research as central, borrowed from the Germans, was given full and acknowledged impetus. A graduate school was established at Columbia University in 1880. Clark University, established in 1889, under the dominating leadership of G. Stanley Hall, soon thereafter offered thorough graduate training in a limited number of fields. In 1892, two years after its establishment, the University of Chicago embarked actively on its career as a leader in university instruction. Shortly before the beginning of the present century three state universities joined the group offering thoroughgoing graduate instruction —the University of Wisconsin in 1892, the University of Nebraska in 1895, and the University of Kansas in 1896.

Educational research, as differentiated from the narrower field of research in educational psychology, had its beginnings before the present century. Only one event in the broad history of educational research is discussed. This is the beginning of the measurement of achievement. Objective achievement measurement in education had its beginning in 1897 with the publication of two papers on achievement in spelling written by Joseph M. Rice (1897), a physician turned educator. The question to which he addressed himself in these papers was whether or not schools spending varying amounts of time in spelling drill showed corresponding difference in skill of their pupils. From his results he concluded that since there was no consistant relation between the achievement exhibited and time spent, the differences he found were due to differences in the quality of teaching. By modern

standards the design he used in the studies was inadequate, but his critics were no better equipped than he in terms of sophistication of methodology. A country-wide controversy sprung up about his criticism of the quality of teaching which lasted for some years thereafter.

Another development related to educational psychology, taking place during these years, was the study of the psychology of learning. A short account of its beginnings is appropriate. It was the work of Ebbinghaus (1913), first published in 1885, which had forcefully called the attention of psychologists to the possibility of controlled research in learning. His work demonstrated that learning could be measured by accurate techniques. Thereafter, research in this area was pursued with enthusiasm. For example, before 1900 Meumann (1913) began research in experimental pedagogy and wrote an influential book which was subsequently translated into English. Despite the fact that the conceptual significance of learning was not fully recognized, as shown, for example, by the failure of Baldwin's *Dictionary of Psychology*, published in 1902, to even list learning as a topic, research activity was not inconsiderable before the new century. In a selected bibliography, published nearly thirty years ago, McGeoch (1933) cited 1200 titles. Six of these papers were published before 1890 and 60 more articles before 1900. Before 1900, investigators had performed studies of the influence on learning and retention of age differences, interference, the curve of learning, reflex "control," motor skills, direction of association, the influence of distraction, the distribution of practice, diurnal variations, fatigue, frequency, habit formation, individual cues, intelligence, knowledge of results, degree of meaningful organization, memory span, racial differences, recitation, sense modality, time interval in associations, transfer, variability and difficulty of material, the whole-part problem, affective memory, childhood memories, context conditions, curve of retention, incidental memory, recognition, reminiscence, and repeated reproduction. Many of these studies were by researchers who did not systematically follow through after their first effort in this area. Content to publish one or two studies, thereafter no sign of continued publication could be observed. Nevertheless, a respectable beginning on problems of learning had been made before the turn of the century.

THE INCUBATION OF EDUCATIONAL PSYCHOLOGY: 1900-1918

If the beginnings of educational psychology took place before 1900, most of the next two decades was its period of incubation. During these years psychologists began to specialize in educational psychology. It was the period of Edward L. Thorndike and Charles H. Judd, the first educational psychologists. It was also a time of intense activity concerning the learning and reading processes, activity in large measure stimulated by these two men. Intelligence and achievement testing also received considerable attention during these years.

Edward L. Thorndike. The first man to deserve to be called an educational psychologist in the modern sense of the term is Edward L. Thorndike. Almost his entire academic life was spent at Teachers College, Columbia University, where for over 40 years he labored on problems of educational psychology. Generations of students, including both prospective teachers and psychologists, came under his influence. He held that educational procedures of all sorts should be based upon results of psychological research, not upon opinions. He stated his position bluntly and succinctly in the preface to a book published three years after the turn of the century. "This book attempts to apply to a number of educational problems the methods of exact science. I have therefore paid no attention to speculative opinions and very little attention to the conclusions of students who present data in so rough and incomplete a form that accurate quantitative treatment is impossible" (1903, p. V).

It was Thorndike who patiently and consistently first systematized the study of learning. Beginning his work in animal learning while a graduate student at Harvard under James, Thorndike performed quantitative experiments on animal learning (Murphy, 1949). The 1897 study of the learning of cats in a puzzle box is a landmark in the history of animal psychology. Although a little later he concentrated on learning in humans, he maintained the same interest shown in this study in that he continued to be concerned with the nature of the learning curve and the conditions which affected it. His early research at Columbia eventuated in his *Elements of Psychology* (1905) which contained his formulation of a variety of "laws of learning"—exercise, effect and readiness.

Thorndike and Woodworth at the turn of the century published their famous paper on transfer of training. Thorndike followed this with a series of studies designed to verify or to refute the doctrines of transfer of training and formal discipline. The general conclusion from these studies was that ideas or habits acquired in one sphere of activity were transferred to another sphere only if there were common elements shared by both spheres.

He was also interested in arriving at some conception of the nature of learning. Learning was seen as forming bonds or connections between stimuli and responses. Hence he combatted the faculty point of view that learning was the training of the potential faculties. Some years later, with the appearance of studies by others on the basis of which insight was postulated as a form of learning, he vigorously disagreed, stressing that the formation of bonds of association was a gradual process even though they sometimes gave the impression of sudden appearance. When in 1913-1914 his three volume edition of *Educational Psychology* came out, most of the experimental material included was derived from his own original investigations. Although his theoretical interpretation of his studies of learning has since been

questioned, no one can deny his great influence upon the field of learning.

Although at the core of Thorndike's efforts were his numerous studies of learning, it is astonishing to find in how many other areas he worked in these early years of the century. In 1902 he published the first work to make generally accessible the statistical approach to the problems of education, sociology, economics and anthropology, *The Theory of Mental and Social Measurements* (1902). In it he considered measures of central tendency, variability, and correlation and so made available to the student an understanding of the treatment of data which was beginning to flow in increasing amounts from the laboratory and classroom. In 1907 he prepared a report for the U.S. Bureau of Education on the elimination of children from the schools who could not profit from the curriculum, a forerunner of many other reports on this and related problems. Thorndike promoted the measurement of ability in school children and developed scales for the measurement of subject-matter fields. It was in 1908 that Stone published his arithmetic test, which is considered to be the first standardized achievement test. Before 1917 Thorndike and his students or associates constructed and standardized a considerable number of other tests. By the name of the investigator, the subject matter field and the date of appearance they were as follows: Curtis, Arithmetic, 1908; Thorndike, Handwriting, 1909; Hillegas, English Composition, 1912; Buckingham, Spelling, 1913; and Thorndike, Reading, 1914, 1916.

Charles H. Judd. Sharing with Thorndike the honor of being a pioneer educational psychologist was Charles H. Judd. He was trained at Leipzig by Wundt, receiving his degree in 1896. His *Genetic Psychology for Teachers* (1903) exerted considerable influence upon school practice. In it he stressed the psychology of the school subjects of reading, writing and arithmetic and forcefully stated the doctrine of biological and psychological development. After some years of other academic appointments, he came to the University of Chicago as Director of the School of Education, a post he held from 1909 to 1938. He took part in many educational surveys and edited various monographs and periodicals, including the *Elementary School Journal* and the *School Review.*

It was in the field of the psychology of reading that Judd's work was particularly outstanding. While at Yale University, he developed a method for the photographic study of eye movements. The studies of reading which were carried out at the University of Chicago were largely under his inspiration or direction.

There had, of course, been earlier work on reading. Perceptual studies of the process of reading had been conducted as early as 1844 by Valentius (Gray, 1960). This was followed by the work of Cattell, Erdmann and Dodge in such problems as the study of eye movements. Details of this early history may be found in Huey (1908).

Gray (1960) has made a tabulation of the number of scientific studies relating to reading published in the United States and England by decades beginning with 1881. In the three decades before 1911 there were 32 such studies. But in the decade 1911-1920 alone there were six times as many studies, with 200 reported by him. According to the same reviewer prior to 1910 most of the research had been concerned with the psychology and physiology of reading. The period from 1911 to 1920 is characterized by him as a transitional one since it marked the beginning of an interest in a broader conception of the field and showed a more clear recognition of the need for applying objective techniques to classroom problems. Moreover, it was during this period that the new tests were introduced, making it possible to study under classroom conditions large groups of children. Further details may be found in Gray (1960).

Intelligence Testing. Utilization of the Binet test in the United States came about particularly through the efforts of H. H. Goddard, F. Kuhlmann, and L. M. Terman, three Ph.D.'s from Clark University. In 1908 Goddard translated the Binet Scale into English and applied it to American children. In 1912 Kuhlmann introduced in the United States a revision of the Binet Test. In 1916 Terman extended and modified the Binet Scale, retaining its good features and eliminating some of its weaknesses, so that it was especially adapted for use in the United States. Although the first two aforementioned forms were used in research, it was Terman's revision, of the so-called Stanford-Binet, that became the indispensible, widely used instrument for testing the intelligence of individual children. To designate the degree of intelligence of the subject, irrespective of age, in 1912 Stern (1914) had advanced the concept of the intelligence quotient. This useful procedure was adopted by Terman in his revision of the test.

By 1917 the literature on the Binet tests had proliferated to the extent that in that year Boardman (1917) was able to report a bibliography of 344 articles and books. Of the writers, H. H. Goddard was most prolific, having published 28 articles concerning the Binet between 1908 and 1916. Binet and Simon, either alone or in collaboration, supplied 19 titles. J. E. W. Wallin published 16 reports. The publications of L. M. Terman, by this time, numbered 15, including one published in 1906 concerned with a specific test from the scale. His first report on the Binet-Simon Scale as a whole was published in 1911.

Achievement Testing. Prophetic of the many later studies of learning in the classroom was the one by Ayres (1909), who studied individual differences in learning through differences in school achievement. As a means of presenting his results he used the method of grade overlapping. His results received wide publicity.

According to Scates (1947), the ten years after Rice's 1897 publication was the period in which the measurement idea incubated. The

work of Thorndike, already mentioned, was important during this decade. In 1912 achievement tests were used for the first time in a large school survey, taking place in New York City. As found by Barr (1960), over 125 similar surveys were conducted in the ten years that followed. Despite this progress, objections to the use of standardized achievement tests were numerous. A favorite device of protagonists for the new tests was the study of the reliability, or, rather, the unreliability, of essay examinations which were the long-entrenched alternative to achievement tests. The studies of Starch and Elliott (1912, 1913a, 1913b) uncovered what was taken to be an amazing lack of agreement among teachers in grading essay-type examinations in the various high school subjects. This and similar evidence resulted in an increasingly favorable reaction to objective tests.

By 1915 enough support for objective testing was available at a meeting of the National Council of Education to bring about a clear-cut decision in its favor. According to Judd (1925), this made such testing respectable in general educational circles. Objectivity and measurement were now established as generally accepted aspects of educational research. It was no accident that it was during these years that Thorndike offered his dictum that everything exists in some amount and that this allowed measurement of that amount.

Other Developments. The National Society for the Study of Education was founded in 1902, but had derived its being from the National Herbart Society founded in 1895. Yearbooks published by the Society include many directly pertinent to educational psychology. They continue to be extremely influential.

The educational research bureau was one of the characteristic developments in the second decade of this century. According to Froelich (1960), the first such bureaus were established in connection with city school systems—Baltimore in 1912 and others less than two years later in Rochester, New Orleans, New York City and six other cities. College and university research bureaus were established in 1913 and 1914 at Oklahoma University, Indiana University, the State University of Iowa, and what is now Kansas State Teachers College (Barr, 1960). By 1925 there were 22 college educational research bureaus, and by 1932 there were 64. State department research bureaus were also organized just before and after World War I. By 1915 there were enough of these bureaus for the National Associations of Directors of Educational Research (now the American Educational Research Association) to be formed (Buckingham, 1941). The *Journal of Educational Psychology* was founded in 1910.

Growth of the Field. Some idea of the growth of educational psychology over the years to about 1918 may be gained from examination of the relative percentage of publications in this field as compared to those found in other areas of education. Franke & David (1931) made

a survey of published educational research, irrespective of field, over the four decades of 1890, 1900, 1910, and 1920. They classified into fields the articles appearing in 13 educational periodicals (*Education, Educational Review, Journal of Genetic Psychology, Teachers College Record*, etc.). In 1890-1899 child psychology had the highest percentage of articles devoted to it, with curriculum, character, educational measurement and statistics coming next in that order. Next to the last, with about 3 per cent of the publications for that decade, was educational psychology (so defined as to refer mostly to studies of learning). The decade 1900-1909 showed similar percentages. By 1910-1919 educational psychology had moved up to be tied for third place with about 10 per cent of the pages of these journals devoted to it. This position was maintained in the following decennial period, 1920-1929. By 1918 educational psychology had been established as an important field of research and its period of incubation drew to a close.

THE RECENT PAST IN EDUCATIONAL PSYCHOLOGY: 1918-1941

A survey of 10 years of educational research covering the period of 1918 to 1927 was published by Walter S. Monroe (1928). His choice of 1918 as beginning a new period of educational research was based on the following considerations: It was during 1917 that psychological testing of Army recruits began, which tended to create a popular interest in the measurement of intelligence; it was in 1917 that the Iowa Child Welfare Research Station was authorized; and it was in 1919 that the Commonwealth Fund, noted for its subventions for educational research, was established, as were the American Council in Education, and the Bureau of Educational Research at the University of Illinois. I would add that it was in 1917 that Ayres (Scates, 1947) stated that measurement had been accepted by the public, and it was in 1919 that, as Buckingham reminded us, test materials were first produced by commercial publishers. Gates and his associates (1948) assert that about 1920 educational psychology could be claimed to have taken definite form. Moreover, a contrast in the treatment of psychology in relation to education became apparent at the beginning of the twenties. This may be illustrated by two widely read authoritative books on the history of education published in 1905 and 1920. One book by Paul Monroe (1905) closes his discussion of psychology with the work of Herbart. The other by Ellwood Cubberly called psychology "the master science" (1920, p. 255), and proceeded to document this claim. For these reasons the recent past in educational psychology is considered to have begun about 1918. Events from the time of World War II onward are too close to hope to see them in historical perspective. Hence, events after about 1941 are not considered.

For the recent past it is impossible to approach the history of educational psychology biographically within the limitations of article

length. Too many men have contributed to this period. Even though some of them are mentioned in the account to follow, perhaps an equal or even greater number of other important contributors have been neglected.

Major Researchers of the Earlier Years. In their survey Monroe and his associates (1928) were able to assemble about 3700 "worthwhile" references to educational research for the decade 1918-1927, the earlier years of the recent past. From among these they selected five major research projects as the most highly important that were published during that period. The first of these was "the Chicago reading studies." Between 1920 and 1925, Judd and his colleagues, Buswell, Terry, and W. S. Gray, published reports of six major studies carried out in the reading laboratory at the University of Chicago. The studies concerned the eye-voice span, the reading of numerals, the development of reading habits, remedial cases in reading, the kinds of silent reading, and, finally, a summary of investigations related to reading. The Thorndike study of the measurement of intelligence in the period 1922-1925 which culminated in the book by Thorndike and his associates, *The Measurement of Intelligence,* was another project selected by Monroe. The studies of genius conducted by Terman and his co-workers, including Katharine Cox, were also selected as among the important studies of that decade. This project was begun in 1921 and the decade saw publication of both Volume I and Volume II of the study. The fourth major project was the nature-nurture studies reported in the *Twenty-seventh Yearbook of the National Society for the Study of Education* (Whipple, 1928). The yearbook was devoted to a large number of papers concerned with the influence of nature and nurture upon intelligence and achievement. The studies dealt with the causes and significance of the large individual differences in intelligence and achievement test results. Among the participants were Barbara Burks, Truman Kelley, E. L. Thorndike, H. E. Jones, F. N. Freeman, K. J. Holzinger, Florence Goodenough, Gertrude Hildreth, Arnold Gesell, Leta Hollingworth, W. McCall and J. Peterson. The papers included one by Burks and Kelley on statistical hazards in nature-nurture study; several on family resemblance in intelligence test scores; others on the relation of intelligence to social environment, to race differences, to schooling, to health, to physique, and to coaching. Still others were devoted to the relation of achievement to intelligence, school attendance, school methods, school expenditures, effort, and mechanical ability. Since potentially any kind of educational research might have been selected, it is significant that it was only the fifth study which did not fall within the scope of educational psychology. This was an investigation, begun in 1921, concerned with educational finance.

Intelligence Testing. As a field of study and application educational psychology received considerable stimulation during the earlier years of the recent past from the development of the intelligence test

arising from the work of Binet and those that came after him. Several lines of evidence attest to the attention to intelligence testing. Although it was opposed at first, acceptance on the part of educators of the concept of intelligence as the mass indicator of intellectual maturation in their students was finally won. Parents, by and large, came to accept application of such tests although perhaps grudgingly. With changing views of the function of the schools, no longer was such measurement seen as an invasion of privacy as it was in the days of Hall when he wished to get height and weight measurement of pupils in the schools. The amount of research appearing before 1917, whch was stimulated by the Binet Test, documents the interest. Even more numerous studies appeared after World War I. The development of the sub-specialty of the mental tester and, indirectly, of the clinical psychologist, was fostered by this interest, since administering the Binet became what was practically a full-time occupation for a considerable number of persons.

The Stanford-Binet was seized upon by educators and psychologists when it was recognized that it met certain already existing needs (Goodenough, 1949). Compulsory school attendance, the increase in the length of school period, and the increasing number of backward students in the schools created a favorable situation for its avid use. These same educational problems also made desirable the development and utilization of instruments for the large scale assessment of groups.

Group Testing. World War I gave special impetus to the interest in group intelligence testing. There had been some preliminary work by Binet and Simon (Peterson, 1925), and selection and the use of group intelligence tests in the U.S. had been begun by Arthur Otis and L. L. Thurstone a few years before the entry of the U.S. in the War. For example, in 1915 the latter had started a group testing program for prospective students at Carnegie Institute of Technology. It was the former, however, whose first work with group tests helped in the development of the Army Alpha and Army Beta which were used in selecting Army recruits. About 1,750,000 men were given these tests. A detailed account is to be found in the report edited by R. M. Yerkes (1921), who was in charge of this work during the War years.

The twenties saw the application of group intelligence tests to school problems. The dates that representative well-known tests were published is of some interest (Garrett and Schneck, 1933). *The Otis Group Intelligence Scale (Advanced Examination)* appeared in 1918. *The Terman Group Test of Mental Ability* and *The Haggerty Intelligence Examination, Delta II* both appeared in 1920. *The Otis Self-Administering Tests of Mental Ability* was published in 1922, and the *American Council Psychological Examination* of the Thurstones first appeared in the 1924 edition. Both the *CAVD Intelligence Scale* of Thorndike and the *Kuhlmann-Anderson Intelligence Tests* appeared in

1927. Group testing of intelligence was established as an important phase of educational psychology by the beginning of the thirties.

Although attention to theoretical formulations was not entirely absent, most psychologists were relatively uninterested in the nature of what they were so busily measuring. In an attempt to remedy this situation, a symposium was held on the meaning of intelligence and published (Thorndike, *et al,* 1921) in the *Journal of Educational Psychology.* The reactions of the participants were diverse and contradictory but the overall effect of the symposium was salutary. S. L. Pressey probably represented the majority of psychologists of that time when he asserted at the symposium that he was more interested in what a test would do than he was in defining intelligence. Nevertheless, it is of some interest to examine briefly some of the participants' remarks. The points stressed were as follows: Thorndike—the goodness of response from the standpoint of truth; Terman—abstract thinking; Calvin—learning to adjust; Pintner—a means of adapting to new situations; V. A. C. Henmon—a combination of intellect and knowledge; Peterson—a mechanism for adjustment; Thurstone—a combination of inhibitive capacity, analytical capacity, and perseverence; Woodrow—the capacity to acquire capacity for performing acts successfully; Dearborn—the capacity to profit by experience; and Haggerty—the activity of a group of complex mental processes. The functional cast of the definitions by these leading educational psychologists is overwhelmingly evident.

Attempts at the measurement of personality lagged behind those concerned with the measurement of intelligence and achievement. The prototype of paper and pencil personality questionnaires was the Woodworth *Personal Data Sheet,* begun during World War I, not used during the war but appearing shortly thereafter. The *Mental Hygiene Inventory* of House appeared in 1926. It was not until 1929 that the *Personality Schedule* of L. L. and T. G. Thurstone was published. In 1931 both the *Personality Inventory* of R. G. Bernreuter and the *Emotional Maturity (E.M.) Scale* of R. R. Willoughby made their appearance.

Another form of psychological testing important to educational psychology, that of the tesing of interests, came into prominence in the 1920's and the 1930's. Earlier Herbart, Dewey, Hall and Thorndike (Fryer, 1931), each in his own way, had emphasized the importance of the factor of interest to the psychology of education. As a consequence, research on the measurement of interest was inevitable. Work was begun at the Carnegie Institute of Technology in 1919 upon the earliest standardized interest inventory in the sense that term is used today, *i.e.,* statistical evaluation of the instrument and the working out of an objective scoring method. Clarence Yoakum, Bruce Moore and J. B. Miner were important figures in this early work. Cowdery (1926-1927) revised the Carnegie Inventory in 1924 and used it for differentiating

groups on the basis of their interests. E. K. Strong took up this form of investigation. The now very well known Strong *Vocational Interest Blank,* which first appeared in 1928, was a thorough revision and extension of the Cowdery Inventory (Fryer, 1931). Many studies using this measure as a research instrument appeared in the thirties.

Under the leadership of Ben Wood, The Cooperative Test Service began in 1932 publishing a considerable number of comparable forms of tests.

Child Development. Hall and others had already established an interest in the growth and development concept. Even before the beginning of the period in question a considerable amount of material was available. One illustration will suffice. In March, 1914, a bibliography of the available representative books in child study was published by Louis N. Wilson (1914), then librarian of Clark University. The first section was devoted to general child psychology and, aside from the usual classic sources of Darwin, Hall and Preyer, 12 other books published since the turn of the century were included. Among the authors cited were Edwin A. Kirkpatrick, Kathleen C. Moore, Milicent Shinn and Edward L. Thorndike. Eleven books or monographs concerning the application of child study to education were listed. The health of the child received eight citations including a book by L. M. Terman, *The Hygiene of the School Child.* Four references to defective children and a similar number of reminiscent studies of childhood followed. The bibliography concluded with mention of the four child study journals then available: *The Child, Child-Study, Journal of Educational Psychology,* and the *Pedagogical Seminary.*

Arnold Gesell, one of Hall's proteges, eventually combined pediatric and psychological interests in his meticulous patient studies of mental and physical development in children (Gesell, 1952). After taking his degree in psychology, he had studied medicine in order to learn more about the physical bases and psychological processes of growth. In 1911 while connected with Yale University, he was assigned space in the New Haven Dispensary for the establishment of a psycho-clinic for children. This was the beginning of the Yale Clinic of Child Development. In 1924 came his first contact with cinema when he prepared a film on the behavior development of the pre-school child from early infancy to school entrance. One of his first major works, *The Mental Growth of the Pre-School Child,* appeared in 1925. This volume was one of a long series devoted to what are now his well known studies of child development. About 1926, the Clinic started to take on its modern form, complete with one way vision screens, research associates and a systematic plan for the recording of infant behavior.

One of the first systematic studies of physical and mental growth in a variety of its manifestations was carried on under B. T. Baldwin, director of the Iowa Child Welfare Station from 1917 to 1928. In

1921 Baldwin published a classic study of the physical growth of children with primary emphasis on anthropometric measures. He worked out the growth curves of these measures including some for individual children, repeatedly measured.

In 1922 W. F. Dearborn started the so-called Harvard Growth Study in which he followed a population of children from the time they entered the first grade until they finished high school (1938, 1941). Both mental and physical tests were used.

The technique of longitudinal study was followed at various other so-called institutes of child welfare which began to function throughout the country during these years. Funds from Laura Spellman Rockefeller and the General Education Board supported this development. In addition to those already mentioned, institutes at the University of Minnesota, the University of California, the Merrill-Palmer School and the Fels Research Institute became active, especially in areas where long continued studies of children of wide scope were necessary. In 1935 the Society for Research in Child Development was organized.

The Influence of Schools of Psychology. Behaviorism, psychoanalysis and the Gestalt psychologies had appeared on the psychological scene before the beginning of the period under discussion. However, some time elapsed before their impact upon educational psychology was perceptible. The onset of their influences occurred in the twenties and thirties.

The militant espousal of an objective approach, the behaviorism of John B. Watson, had made its appearance beginning in 1912, culminating in his 1919 publication, *Psychology from the Standpoint of a Behaviorist.* It took some years for its effect to be seen in educational circles.

What is the nature and extent of the effect of behaviorism upon educational psychology? Butts and Cremin (1953) consider behaviorism to have performed the function of serving as a more extreme version of associationism. They do not go on to cite specific instances of its effect upon education. In fact, here or elsewhere, specific influences of behaviorism upon educational psychology are not too evident. An illustration drawn from a review of behaviorism is relevant. In 1938 Harrell and Harrison (1938) published a paper entitled, "The Rise and Fall of Behaviorism." They cite 426 references. These references were searched for items of educational import. They mention incidentally that J. S. Gray (1935) wrote a manual of educational psychology from the standpoint of a strict objectivist. This is their only reference to educational psychology. Turning to this manual, in the preface Gray made it clear that he adheres to no specific school, and that he wished to stress objective data, particularly what he called, "biological data," leading him to stress such work as that of C. J. Herrick, G. E. Coghill,

and C. M. Child. No evidence is available to the writer about the reception of the book. Others similar in spirit and addressed to problems in educational psychology do not seem to have been published. To return to Harrell and Harrison, search of their bibliography reveals one reference in the *Journal of Educational Psychology,* one in the *Hibbert Journal* and one in *School and Society.* Searching their contents showed that one had but an incidental mention and both of the others were critical of behaviorism. Of course, I am confident that if a patient, complete search was made of educational periodicals other articles referring to behaviorism would undoubtedly be found. However, no article of this sort seemed to have reached the stature of being considered as a major contribution to educational psychology so far as the writer is aware. One is left with the impression that behaviorism during these years (or later) did not have much specific effect upon educational psychology.

The Gestalt psychologies arising from the work of Wertheimer, Köhler, Koffka, and Lewin came into educational psychology in a variety of ways. The publication of the book by R. M. Ogden (1926), *Psychology in Education,* called attention to the value of Gestalt psychology to students of educational psychology. In 1931 Seltsam was stimulated to report the main contentions of Gestalt psychology in the *Journal of Educational Psychology,* because of what he said was an increasing interest and concern with this approach for which he sited specific instances. The effect of Gestalt psychology upon education has helped to give an integrationist view of human behavior. The influence of Gestalt psychology was probably enhanced by its affinity with the independently developed progressive education movement. As Carr (1934) indicated, Gestalt psychology and progressive education have many points of compatibility, if not identity. He suggested as illustration, the Dewey-Kirkpatrick principles of total rather than isolated learning, and their premise that parts of learning should be learned in relationships to the whole to which they belong. A number of parallels are dealt with in detail. Not only has Gestalt psychology integrated to an extent with progressive education, but it has also served a function, in the view of many workers, as a corrective to the extremes of connectionism and behaviorism (*e. g.,* Butts & Cremin, 1953). The influence of the work of Lewin upon social psychology and, indirectly, upon social psychology of childhood is self-evident.

It is probable that no one would deny that psychoanalysis had an influence upon educational psychology. As was the case with behaviorism, it is difficult to be specific about this matter. The problem is complicated by the fact that its influence has probably been indirect in that some other vehicle such as child guidance clinics, the progressive education movement, mental hygiene and the like, were first influenced, and they in turn influenced educational psychology. That it has had an influence is attested to by the study performed by Park (1931). She

analyzed the contents of general psychology texts and found that not
a single one of the books examined published between 1917 and 1930
failed to show, either explicitly or implicitly, the influence of psycho-
analytic doctrines. She concluded that this influence increased steadily
between 1910 and 1930. It is plausible to believe that psychoanalysis
had an influence upon guiding the thinking of educational psychologists
in the direction of emphasis on the importance of the early years,
stress on the irrational facets of the child's makeup, on the necessity
of warmth and permissiveness and, more generally, on increased at-
tention to all facets of the personality and makeup of the child rather
than to his mind or intellect alone.

Courses in Educational Psychology. Using 1927-1930 as the base
years, Gates (1932) analyzed various studies concerned with the pro-
grams of courses in educational psychology in representative teacher
training institutions. By and large, these studies had all found a con-
siderable increase in the total number of courses over those offered
during the previous decade. As Gates indicates, the expansion con-
sisted mainly of developing as separate courses some of the topics
formerly treated more briefly in the general courses in educational
psychology. He also attributed the increase to the fact that there has
been a lengthening of teacher training with the older types of two-
year normal schools increasing their offerings to three or four years.
Meanwhile, the liberal arts colleges had been adding to their curricula
a sufficient number of courses in education to meet the requirements
for a teachers certificate. Up to this time this had been a relatively in-
frequent practice.

The Concept of Evaluation. It was during the thirties that Ralph
W. Tyler of Ohio State University and his associates gave "a new di-
rection" to measurement through the concept of evaluation (Smith &
Tyler, 1942). The term refers to measurement of student progress, not
so much in terms of achievement as the term is ordinarily used, but
rather in terms of the broader social objectives of education. This
means that a value scale is substituted for a measurement scale. Meas-
urement is a matter of "moreness." Value has to do with appropriateness,
which falls on an optimum point along a scale with the rest of the scale
extending on both sides of this point (Scates, 1947). In this perspec-
tive, amounts beyond this optimum may be as damaging as having too
little. The nature of evaluation and some of the cautions in its con-
ceptual application are given in Lorge (1941).

Progressive Education. In a report published in 1941, a commit-
tee of the Progressive Education Association (Baker, 1941) reported an
analysis and summary of recent comparative research studies of old
and new methods in education. The committee that wrote this report
was chaired by G. Durwood Baker, a superintendent of schools, and
had as some of its members R. Travers, A. Eurich, I. Lorge, P. J. Rulon,

R. Tyler, Elizabeth Woods and J. W. Wrightstone. In the preface they state that it was only during the '30's that public school systems throughout the country attacked the problem of revising and modernizing their courses of study. This volume is an important summarization of what has been referred to as progressive education. They define progressiveism as carrying out a great deal of the learning going on in the school through activities at the expense of a smaller amount of time spent in routine drill. They also emphasized the importance of training for a living, learning by doing, and educating for participation in society. All these are to be carried through by an integrated curriculum, *i.e.,* drawing upon any field of knowledge that will assist the child in fulfilling these objectives.

After reporting a variety of surveys and studies comparing and contrasting older and newer methods of teaching they reached the conclusions that the newer methods do not result in the loss of academic proficiency in the usual school subjects and that, invariably, there is a definite gain in such characteristics as initiative, skill in dealing with problems, and knowledge of world affairs and social participation. They also point out, as an implication of the evidence, that each child has an individuality of his own and that learning in the school must be continuous and not a thing apart.

School Achievement. W. S. Monroe (1945) in contrasting the measurement of school achievement of 1945 with that of 1920 points to the sheer increase in number and kinds of tests, but stresses the ingenuity shown in devising new types of objective exercises, greater attention to the compatability of the exercises and the abilities one is attempting to measure, and greater attention and complexity of item analysis.

Reading. According to Gray (1960), since 1920 the scope of studies of reading have broadened considerably. Since that time recognition has come that this field includes problems that arise before the children begin to actually read and that studies of adult readers are also necessary and relevant. Increased appreciation of the complexity in the problem of reading is also apparent.

Learning. In 1940 Stroud published a careful selective review of studies of learning in the school situation. Despite this selectivity he referred to 256 titles. He closed the review with general evaluative statements which are of interest. He concluded: (1) that the discrepancies between the laboratory and school setting is not a serious obstacle to interpretation; and (2) that the research in the school setting compares not too unfavorably with that done in the laboratory, although not as good as the very best done in the laboratory. He goes on to a plea for closer cooperation between experimental and educational psychologists. Munn (1942), reached a somewhat different conclusion after reviewing the classroom application of learning. Al-

though there are leads, he feels that in most instances application to classroom conditions and curricular materials is yet to be worked out. He sketched the two extremes of the positions taken by various educational psychologists. On one hand, there are those who would transfer practically any kind of results of the study of learning to the classroom; at the other, there are those who consider relevant only those studies performed in the classroom itself. He feels the truth lies between the extremes.

The 1942 or Forty-first *Yearbook of the National Society for the Study of Education* (Brubacher, 1942, Henry, 1942) is a landmark in the history of educational psychology. It was devoted to the psychology of learning. The first of the two sections was concerned with theories of learning. Guthrie, Hull and Lewin presented their own views of learning while connectionism was represented by articles by Gates and by Sandford, and field theory by Hartmann as well as by Lewin. An attempt at a reconciliation of the learning theories by McConnell closed this section of the *Yearbook*. He finds in their accounts points of synthesis and reconciliation. In fact, he lists nine points of similarity. He is commendably cautious in his interpretation, hoping that the consideration of the theories will stimulate further research. The second major section of the report (Henry, 1942) was concerned with the implications for education of the various viewpoints expressed in the first section. Ryans discussed the implications for motivation; J. E. Anderson those for emotional behavior; Stroud those for practice; Horn those for language and meaning; Brownell those for problem solving; and Buswell those for the organization and reference to the curriculum.

Social Psychology. It is prophetic that in 1941 Trow called attention to the fact that most educational psychologists had been content to limit themselves to the psychology of the individual, while neglecting the social facets of his behavior. The pressures for attention to social factors came not so much from educational psychologists as from so-called progressive educators, curriculum specialists, youth leaders and the like. Dewey, was, of course, a prime force in this plea for the child to be seen as a social being.

The studies in social psychology relevant to education seem to have been published toward the end of the period of the recent past. To be sure, there had been earlier research descriptive of social development in the normative sense but little research on how this development took place or on the dynamics of the social forces impinging on the child. In the years between 1927 and 1939 Lehmann and Witty (1927) published on play activities, Maller (1929) gave his report on cooperation and competition, Goodenough (1931) published her extensive study of anger in young children, Isaacs' book on social development (1933) was published, the study of childhood conflicts by Jersild and Markey (1935) appeared, Lois Murphy (1937) published

her book on social behavior and personality, Lewin, Lippitt, and White (1939) published a major paper on "social climates," and in the same year Dollard and his associates (1939) published their book on frustration and aggression. These developments in social psychology were important pioneer steps leading to the present day concern with the social aspects of educational psychology. Before these and similar studies made their appearance, educational psychology was primarily oriented toward increasing the efficiency of the learning of the formal subject matters of education, as the preceding discussion shows.

By 1941, the beginning of the extended present, this orientation is seen only as a point of departure since learning takes place everywhere and in all situations whether scholastic or not. The task of the school had broadened beyond imparting formal knowledge and skills. The doctrine of the "whole child" brought in its wake concern over the social and personal development of the child. This state of affairs broadened the scope of educational psychology, but it also tended to blur the nature and extent of educational psychology. Guiding development in the schools had become increasingly complex.

REFLECTIONS ON THE HISTORY OF EDUCATIONAL PSYCHOLOGY

The comments that close this brief survey of educational psychology are confined to some reflections on two of the major trends exhibited.

Another paper (Watson, to be published), concerned with the present status of educational psychology and of educational psychologists, is relevant in considering the significance of the facts that have been brought forth in this paper. In this other paper the nature of educational psychology is shown by presenting descriptions of its scope by various authorities, by examining available information about undergraduate course offerings, by comparing the content of textbooks in the field, by analyzing contributions to it from various other areas of psychology, and by discussing its reciprocal contributions to these areas of psychology. In the paper in question, there is also consideration of educational psychologists in terms of their number, their rate of growth, the positions they hold, the training they receive, the research they carry on, the values they hold, and the prestige afforded them by their colleagues. Thereafter, the various steps being taken to remedy some of the weaknesses that are recognized as existing in educational psychology and among educational psychologists are surveyed. It is significant that the major finding about the present status of educational psychology is widespread dissatisfaction expressed both by psychologists within the field and those outside of it. More specifically, educational psychology is found to have diminished in prestige; research in the field is felt to be uneven, and, on the whole, disappointing; feelings of estrangement are evidenced between psychologists in general and

educational psychologists, and a field is noted that while broadening in scope is losing its former co-ordination. Can we see from this historical survey any of the roots of these sources of dissatisfaction? If one stands back from the mass of historical details that were presented and tries to see the broad outline, the reviewer suggests that two trends may be discerned.

First, there has been an overwhelming emphasis in educational psychology upon the practical application of psychology. Theoretical issues have been neglected. Unless it is argued this reviewer has been guilty of selective projection in writing this account so as to overstress the applied aspect and minimize the theoretical, this conclusion seems self-evident. There is a relative lack of theoretical orientation in educational psychology. Throughout the time span of the historical survey the research problems appear to be selected primarily because of pragmatic interests. Perhaps not since Dewey has there been a strong theoretical guideline for the conduct of research in educational psychology. In a sense the connectionism of Thorndike was a militant espousal of the lack of necessity for a theoretical orientation. Behaviorism, psychoanalysis and Gestalt theory, although they made some impression on educational psychology, arose outside of the field and did nothing more than modify it in some respects. In no way can it be said that any or all of them did anything to revolutionize the field. The theoretical orientations that guide modern approaches to learning seem to have made astonishingly little impression upon research or application of the psychology of learning to the problems of educational psychology. They are not the foundation upon which educational psychologists build their work.

Second, the field of educational psychology has become more complex as the vision of what it encompasses has broadened. Originally concerned with learning and measurement, its scope has been extended with each succeeding generation, to the point where now the newest extension is social educational or educational social psychology. When one is reminded that this reviewer quite deliberately omitted consideration of the professional service functions, such as those in clinical and counseling activities which are also broadening the field, and confined himself to teaching and research, this sense of a broadened field is heightened. As educational psychology becomes broader, it continues to train its own specialists in whatever is the particular new area in question. Consequently, there is more and more separation of educational psychologists from the rest of their brethren. Most often trained in different schools, on different subject matter, and taking jobs in separate places of employment, they continue to lose contact with the rest of the psychological fraternity. At least in these respects the history of educational psychology shows that the field has not as yet realized its full potential.

REFERENCES

ADLER, M. (Ed.) *The great ideas: a syntopicon of the great books of the western world.* Chicago: Encyclopedia Britannica, 1952.

ALLPORT, G. W. The productive paradoxes of William James. *Psychol. Rev.,* 1943, 50, 95-120.

AQUINAS, T. *Summa theologica.* New York: Benziger, 1911-1920.

AQUINAS, T. *Concerning the teacher.* (Quaestiones Disputatae de Veritate, Q. XI) In J. V. McGlynn (Trans.), *St. Thomas Aquinas, the teacher, the mind.* Chicago: Regnery, 1953.

ASMUS, W. Herbart's psychologie und pädagogik des charakters. *Psychol. Beit.,* 1957, 3, 390-406.

AYRES, L. P. *Laggards in our schools.* New York: Russell Sage Foundation, 1909.

BAKER, G. D. (Ed.) *New methods vs. old in American education: an analysis and summary of recent comparative studies by the informal committee appointed by the Progressive Education Association to report on evaluation of newer practices in education.* New York: Bureau of Publications, Teachers College, Columbia Univer., 1941.

BALDWIN, B. T. Physical growth of children from birth to maturity. *Univer. Iowa Stud. Child Welf.,* 1921, 1, No. 1, 1-411.

BARR, A. S. Research methods. In C. W. Harris (Ed.), *Encyclopedia of educational research.* (3rd ed.) New York: Macmillan, 1960, Pp. 1160-1166.

BINET A. & HENRI, V. La memoiré des mots. *Année psychol.;* 1895, 1, 1-23.

BINET, A., & HENRI, V. La psychologie individuelle. *Année psychol.,* 1896, 2, 411-465.

BINET, A., & SIMON, T. Methodes nouvelle pour le diagnostic du niveau intellectual des anormaux, *Année psychol.,* 1905, 11, 191-244.

BLAIR, G. M. *Educational psychology, its development and present status.* Urbana, Ill.: Bureau of Research & Service, College of Education, 1948.

BOARDMAN, HELEN *Psychological tests: a bibliography.* New York: Bureau of Educational Experiments, 1917.

BORING, E. G. *A history of experimental psychology* (2nd ed.) New York: Appleton, Century, Crofts, 1950.

BORING, MOLLIE D., & BORING, E. G. Masters and pupils among the American psychologists. *Amer. J. Psychol.,* 1948, 61, 527-534.

BRADBURY, DOROTHY E. The contributions of the child study movement to child psychology. *Psychol. Bull.* 1937, 34, 21-38.

BRETT, G. S. *A history of psychology.* Vol. II, *Medieval and early modern period.* New York: Macmillan, 1921.

BRUBACHER, J. S. (Ed.) The psychology of learning. *Yearb, nat. Soc. Stud. Educ.* 1942, 41, Part I.

BUCKINGHAM, B. K. Our first twenty-five years. *Addr. & Proc. nat. Educ. Ass.,* 1941, 79, 347-363.

BURT, C. Impact of psychology upon education. *Yearb. Ed.,* 1957, 163-180.

BUTTS, R. F., & CREMIN, L. A. *A history of education in American culture.* New York: Holt, 1953.

CARR, J. W. JR. The relationships between the theories of Gestalt psychology and those of a progressive science of education. *J. educ. Psychol.*, 1934, 25, 192-202.

CATTELL, J. M. Mental tests and measurements. *Mind*, 1890, 15, 373-381.

COWDERY, K. M. Measurement of professional attitude differences beween lawyers, physicians and engineers. *J. Personnel Res.*, 1926-1927, 5, 131-141.

CRABB, A. L. A study in the nomenclature and mechanics employed in catalogue presentation of courses in education. *Peabody Cont. Educ.*, 1926, No. 21.

CRISSMAN, P. The psychology of John Dewey. *Psychol. Rev.*, 1942, 49, 441-462.

CUBBERLEY, E. P. *History of education.* Boston: Houghton Mifflin, 1920.

DAY, J. Original papers in relation to a course of liberal education. *Amer. J. Sci. Arts*, 1829, 15, 300-309.

DEARBORN, W. F., & ROTHNEY, J. W. *Predicting the child's development.* Cambridge, Mass.: Sci-art Publishers, 1941.

DEARBORN, W. F., ROTHNEY, J. W., & SHUTTLEWORTH, F. K. Data on the growth of public school children. *Monogr. Soc. Res. Child Develpm.*, 1938, 3, 1-136.

DENNIS, W. Historical beginnings of child psychology. *Psychol. Bull.*, 1949, 46, 224-235.

DEWEY, J. The reflex arc concept in psychology. *Psychol. Rev.*, 1896, 3, 357-370.

DEWEY, J. *How we think.* New York: Heath, 1933.

DOLLARD, J., DOOB, L. W., MILLER, N. E., MOWRER, O. H., & SEARS, R. R. *Frustration and aggression.* New Haven: Yale Univer. Press, 1939.

EBBINGHAUS, H. *Memory.* New York: Teachers College, 1913.

EBY, F., & ARROWOOD, C. F. *The development of modern education: in theory organization and practice.* (2nd ed.) New York: Prentice-Hall, 1952.

FRANKE, P. R., & DAVIS, R. A. Changing tendencies in educational research. *J. educ. Res.*, 1931, 23, 133-145.

FREEMAN, KATHLEEN. *The pre-Socratic philosophers.* (3rd. ed.) Oxford: Blackwell, 1953.

FROELICH, G. J. Research bureaus. In C. W. Harris (Ed.), *Encyclopedia of educational research.* (3rd. ed.) New York: Macmillan, 1960, Pp. 1155-1160.

FRYER, D. H. *The measurement of interests in relation to human adjustment.* New York: Holt, 1931.

GALTON, F. Psychometric experiments. *Brain*, 1879, 2, 149-162.

GALTON, F. Measurements of character. *Fortnightly Rev.*, 1884, 36, 179-185.

GARRETT, H. E., & SCHNECK, M. R. *Psychological tests, methods and results.* New York: Harper, 1933.

GATES, A. I. The place of educational psychology in the curriculum for the education of teachers. In S. A. Curtis (Ed.), The direct contribution of

educational psychology to teacher training. *Yearb. nat. Soc. Coll. Teach. Educ.* 1932, 20, Pp.21-35.

GATES, A. I., JERSILD, A. T., McCONNELL, T. R., & CHALLMAN, R. C. *Educational psychology.* New York: Macmillan, 1948.

GESELL, A. *The mental growth of the pre-school child.* New York: Macmillan, 1925.

GESELL, A. Arnold Gesell In E. G. Boring, *et al.* (Eds.), *A history of psychology in autobiography.* Vol. IV. Worcester, Mass.: Clark Univer. Press, 1952, Pp. 123-142.

GODDARD, H. H. The Binet and Simon tests of intellectual capacity. *Training Sch.,* 1908, 5, 3-9.

GOODENOUGH, FLORENCE L. Anger in young children. *Univer. Minn. Inst. Child Welf. Monogr. Ser.,* 1931, No. 9.

GOODENOUGH, FLORENCE L. *Mental testing: its history, principles and applications.* New York: Rinehart, 1949.

GRAY, J. S. *Psychological foundation of education.* New York: American Book, 1935.

GRAY, W. S. Reading. In C. W. Harris (Ed.), *Encyclopedia of educational research.* (3rd. ed.) New York: Macmillan, 1960, Pp. 1086-1088.

HALL, G. S. Contents of children's minds. *Princeton Rev.,* 1883, 11, 272-294.

HARRELL, W., & HARRISON, R. The rise and fall of behaviorism. *J. gen. Psychol.,* 1938, 18, 367-421.

HENRY, N. B. (Ed.) The psychology of learning. *Yearb. nat. Soc. Stud. Educ.,* 1942, 41, Part II.

HUEY, E B. *The psychology and pedagogy of reading.* New York: Macmillan, 1908.

ISAACS, SUSAN. *Social development in young children: a study of beginnings.* New York: Harcourt Brace, 1933.

JERSILD, A. T., & MARKEY, FRANCES V. Conflicts between preschool children. *Child Develpm. Monogr.,* 1935, No. 21.

JUDD, C. H. *Genetic psychology for teachers.* New York: Appleton, 1903.

JUDD, C. H. The curriculum: a paramount issue. *Add. & Proc. nat. Educ. Ass.,* 1925, 63, 805-811.

KRUGMAN, M. Orthopsychiatry and education. In L. G. Lowrey & Victoria Sloane (Eds.), *Orthopsychiatry, 1923-1948: retrospect and prospect.* New York: American Orthopsychiatric Association. 1948, Pp. 248-262.

KUHLMANN, F. Revision of the Binet-Simon system for measuring the intelligence of children. *J. Psychol. Asthen. Monogr. Suppl.,* 1912, No. 1.

LEHMAN, H. C., & WITTY, P. A. *The psychology of play activities.* New York: A. S. Barnes, 1927.

LEWIN, K., LIPPITT, R., & WHITE, R. K. Patterns of aggressive behavior in experimentally created "social climates." *J. soc. Psychol.,* 1939, 10, 271-299.

LITTLE. J. K. Graduate education. In C. W. Harris (Ed.), *Encyclopedia of educational research.* (3rd ed.) New York: Macmillan, 1960, Pp. 593-602.

LORGE, I. Evaluation: the new stress on measurement. *Teach. Coll. Rec.,* 1941, 42, 667-679.

McGEOCH, J. A. The psychology of human learning: a bibliography. *Psychol. Bull.*, 1933, 30, 1-62.

MALLER, J. B. Cooperation and competition: an experimental study of motivation. *Teach. Coll. Cont. Educ.*, 1929, No. 384.

MEUMANN, E. *The psychology of learning.* New York: Appleton, 1913.

MONROE, P. *A textbook in the history of education.* New York: Macmillan. 1905.

MONROE, W. S. Educational measurement in 1920 and in 1945. *J. educ. Res.*, 1945, 38, 334-340.

MONROE, W. S., *et al. Ten years of educational research,* 1918-1927. Urbana, Ill.: Univer. of Illinois Press, 1928.

MORRIS, N. L. *William James: the message of a modern mind.* New York: Scribners, 1950.

MUNN, N. L. The psychology of learning and its classroom application. *Peabody J. Educ.*, 1942, 19, 257-265.

MURPHY, G. *Historical introduction to modern psychology.* (Rev. ed.) New York: Harcourt Brace, 1949.

MURPHY, LOIS B. *Social behavior and child personality.* New York: Columbia Univer. Press, 1937.

OGDEN, R. M. *Psychology and education.* London: Routledge, 1926.

PARK, DOROTHY G. Freudian influence on academic psychology. *Psychol. Rev.*, 1931, 38, 73-85.

PETERSON, J. *Early conceptions and tests of intelligence.* Yonkers: World Book, 1925.

RICE, J. M. The futility of the spelling grind. *Forum,* 1897, 23, 163-172, 409-419.

ROBACK, A. A. *History of American psychology.* New York: Library Publishers, 1952.

SCATES, D. E. Fifty years of objective measurement and research in education. *J. educ. Res.*, 1947, 41, 241-264.

SELTSAM, K. Organismic psychology and educational theory. *J. educ. Psychol.*, 1931, 22, 351-359.

SHARP, STELLA E. Individual psychology: a study in psychological method. *Amer. J. Psychol.*, 1899, 10, 329-391.

SMITH, E. R., TYLER, R. W., and EVALUATION STAFF. *Appraising and recording student progress: adventures in American education.* Vol. III. New York: Harper, 1942.

STARCH, D., & ELLIOTT, E. C. Reliability of grading high school work in English. *Sch. Rev.*, 1912, 20, 442-457.

STARCH, D., & ELLIOTT, E. C. Reliability of grading high school work in mathematics. *Sch. Rev.*, 1913, 21, 254-259. (a)

STARCH, D., & ELLIOTT, E. C. Reliability of grading high school work in history. *Sch. Rev.*, 1913, 21, 676-681. (b)

STERN, W. *Psychological methods of testing intelligence.* Baltimore: Warwick & York, 1914.

STROUD, J. S. Experiments on learning in school situations. *Psychol. Bull.*, 1940, 37, 777-807.

TERMAN, L. M. *The measurement of intelligence.* Boston: Houghton Mifflin, 1916.

THOMAS, M. H., & SCHNEIDER, H. W. *A bibliography of John Dewey.* New York: Columbia Univer. Press. 1929.

THORNDIKE, E. L. *An introduction to the theory of mental and social measurements.* New York: Teachers College, Columbia Univer., 1902.

THORNDIKE, E. L. *Educational psychology.* New York: Lemcke & Buchner, 1903.

THORNDIKE, E. L. *Educational psychology.* New York: A. G. Seiler, 1905.

THORNDIKE, E. L. *Educational psychology.* (Three vols.) New York: Teachers College, Columbia Univer., 1913-1914.

THORNDIKE, E. L. *et al.* Intelligence and its measurement. *J. educ. Psychol.*, 1921, 12, 123-147, 195-216.

THORNDIKE, E. L., & WOODWORTH, R. S. The influence of improvement in the mental function upon the efficiency of other functions. *Psychol. Rev.*, 1901, 8, 247-261, 384-395, 553-564.

TROW, W. C. Educational psychology individual or social? *J. consult. Psychol.*, 1941, 5, 265-269.

WALKER, HELEN M. The contributions of Karl Pearson. *J. Amer. Statist. Ass.*, 1958, 53, 11-22.

WARREN, H. C. *A history of association psychology.* New York: Scribners, 1921.

WATSON, R. I. A brief history of clinical psychology. *Psychol. Bull.*, 1953, 50, 321-346.

WATSON, R. I. The present status of educational psychology and educational psychologists. To be published.

WHIPPLE, G. M. (Ed.) Nature and nuture: I. Their influence upon intelligence; II. Their influence upon achievement. *Yearb. nat. soc. Stud. Educ.*, 1928, 27, Parts I & II.

WILSON, L. N. Representative books in child study. *Publ. Clark Univer. Libr.*, 1914, 3, No. 6.

WISSLER, C. The correlation of mental and physical tests. *Psychol. Monogr.*, 1901, 3, No. 6.

YERKES, R. M. (Ed.) Psychological examining in the United States Army. *Mem. nat. Acad. Sci.*, 1921, 15, 1-890.

INFORMATION-GATHERING IN DIAGNOSTIC PROBLEM-SOLVING: A PRELIMINARY REPORT[1]

EDITH D. NEIMARK
New York University

The purpose of the experimental program outlined below is to determine the effect upon diagnostic problem solving behavior of various aspects of the problem situation itself. Four criteria dictated the development of a task for this purpose: (a) aspects of the problem itself (e.g information available, information required for solution, etc.) should be capable of specification and manipulation in a quantitative fashion; (b) just as much of the S's pre-solution behavior as possible should be directly observable and measurable; (c) the task should not required special knowledge or skills; (d) the task should maintain S interest through a series of comparable problems. The task which emerged is formally similar to a test developed by Glaser, Damrin and Gardner (1954) and to the logic tasks developed by a number of investigators (e.g. John and Miller, 1957; Moore and Anderson, 1954). Our concern differs from that of previous investigators in greater emphasis upon task variables per se and in greater concern with observing the development of the solution process than its end product.

Experimental task. The S is presented with a problem board containing m movable shutters and an answer sheet containing a patterns each composed of m binary elements (black or white circles). A simple example for which $m=5$ and $a=4$ appears in Fig. 1. The S is told that one of the patterns on the answer sheet is presented in his board and that if he opened all the shutters he could see it, but that the object of the experiment is to determine which of the patterns is in the board by opening as few shutters as possible. Each time S opens a shutter he is to write the letter designating the position of the shutter on the appropriate line of his answer sheet and across the face of all alternatives which are eliminated as possible answers by the information received from moving that shutter. When S has eliminated all but one pattern he writes the number of that pattern in the line labelled "answer," checks his answer, and proceeds to the next problem. The board contains space for loading as many as eight consecutive problems.

For the majority of studies reported below Air Force recruits served
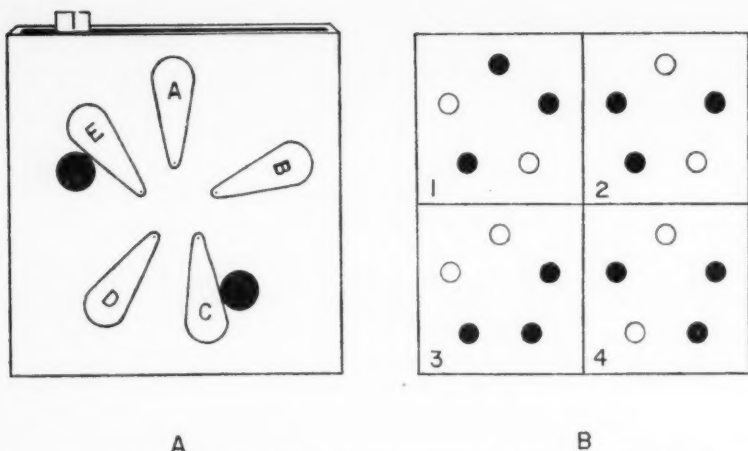
A   B

Figure 1. A. Schematic of a problem board containing five movable shutters with shutters E and C open. B. Array of four patterns which might be presented in the problem board. Only pattern 4 can be the answer to the problem in the board at A.

as Ss; more recently college students have been used. Although the two S populations are comparable with respect to range of age and intelligence, the college population contains girls as well as boys. To date, however, we have found no statistically significant sex differences.

Response measures. Many facets of S's problem solving behavior are directly measurable in this experimental situation. The most obvious measure, (a) number of correct answers, turns out to be an extremely insensitive one since most Ss get all their problems correct. Moreover, as is usually the case, number of correct answers tells very little above how S came up with the particular answer. (b) Total time to solution tends to be lawfully related to a number of variables but, again, it throws little light on the underlying thought processes. (c) Number of shutters opened per problem (which will be called number of moves hereafter) provides a simple gross measure of the amount of information gathering activity preceding solution. The variables to which this measure is related will be discussed below. The best measure of a S's information gathering behavior is provided by (d) the mean expected informational outcome of S's moves on a problem. This measure will be called a strategy score for the sake of brevity; the name is not meant to imply any unobservable planning behavior on the part of S. The strategy score is determined by taking the informational value of each of the two possible outcomes of a move, weighting by the relative frequency of each outcome on the answer sheet, and summing the weighted values. For example, if S chose shutter A of Fig. 1 he would get two bits of information if a black circle appeared and only .415 bits if a white circle appeared. The circle in position A is black on only one of the patterns and white on the other three. Thus, the expected informational outcome of shutter

A as a first move will be .75 (.415)+.25(2.0)=.811. This value is directly obtainable from tabled values of $-p \log_2 p$. The S's strategy score for a problem is the mean of the individual move values computed in this fashion.

Although the strategy score is generally negatively correlated with number of moves (Glaser and Schwarz, 1954), it does provide some unique information. The maximum attainable strategy s c o r e is 1.0, which occurs only when S chooses a series of moves such that each move halves the remaining number of possible answers. Except under unusual circumstances, this strategy should generally be the most efficient one to adopt since it maximizes the expected informational yield (Goldbeck, Bernstein, Hillix, and Marx, 1957). Thus, strategy scores approaching unity reflect rational information gathering; low strategy scores indicate gambling on getting a lot of information quickly.

The fifth measure is a measure of S's efficiency in utilizing the information he has obtained. It is gotten by comparing the information available from a move with the information S actually obtains (as inferred from alternatives bearing the letter of a shutter position across their face) and is expressed as the ratio of the two values. At the present time data on information utilization have not been analyzed in detail.

Independent variables. The response measures listed above are being studied in relation to three broad categories of independent variables: (*a*) informational aspects of the problem; (*b*) reinforcement and training conditions; (*c*) information costs. Under the first heading one obvious variable is the total amount of information needed, which may be measured in bits of information. The amount of information needed for solution is a linear function of $\log_2$ of the number of alternative solution patterns. Results of a study comparing solution behavior of independent groups given arrays containing 8, 16, or 32 patterns show that both number of moves to solution and total time to solution are increasing linear functions of $\log_2$ number of patterns, *a*. Strategy scores, on the other hand, are independent of *a*. Thus, the more information S needs for solution the more he collects, but the nature of the collecting procedure is independent of amount. Furthermore, the relation between amount of information needed and number of information-getting moves is predictable from information theoretic considerations.

Information gathering behavior would *a priori* seem to be a function not only of the amount of information needed but also of the amount and kind of information available. The amount and kind of information available to S in the present study is determined by (*a*) the patterns of the answer sheet, and (*b*) the particular problem presented in the board (the correct answer). Each of these two information

sources, in turn, has been quantitatively characterized and manipulated as an independent variable.

The array of patterns on an answer sheet is characterized in terms of the frequency distribution of expected informational outcome, E_1, of each of the m possible moves as a first move. For example, considering each of the five possible moves of the four patterns in Fig. 1 as a first move, two moves (C and E) have an expected informational outcome of one bit; two (A and D) have an expected informational outcome of .811 bits; and one (B) has an expected informational outcome of 0 bits. The mean of this frequency distribution is $E_1 = .724$ bits. In all studies to date the mean and variance of the E_1 distribution have been held approximately constant and the form of the distribution has been varied: rectangular, unimodal, or multimodal. Although both moves and strategy scores appear to be affected by the form of the E_1 distribution, strategies seem to be the more directly affected measure. In general, bimodal E_1 frequency distributions containing only two widely different E_1 values (rational or gambling moves) seem to make for more rapid adoption of appropriate information gathering strategies. The Ss are most confused when confronted with (a) a large heterogenous distribution of E_1 values, or (b) a homogenous distribution of $E_1 = 1$ (i.e. information maximizing moves). The latter finding was reported earlier by Goldbeck, et al (1957).

The information associated with each specific pattern is also characterized by a frequency distribution, in this case, of information which would actually be obtained from each move as a first move, I_1. The mean informational outcome for a pattern, I, was systematically manipulated for a variety of E_1 distributions.[2] The results show an inverse relationship between I and number of moves; i.e. patterns containing more informative elements are solved in fewer moves. If one assumed that Ss chose moves at random until a problem is solved, then number of moves, N, should be a hyperbolic function of I. On the other hand, if one assumes that Ss always adopt an information maximizing strategy, then N would be independent of I—which it clearly is not. The failure of both models to describe the empirical function suggested by a sizeable number of data points suggests the non-startling conclusion that Ss are neither completely logical nor completely random in their information gathering behavior. Strategy scores, by the way, bear no apparent systematic relation to I (except to the extent that I is involved in the manipulation of reinforcement schedules).

At a more practical level the finding of a relationship between N and I provides an *a priori* basis for the development of problem sets of comparable difficulty; for a given E_1 distribution problems of equal I value are solved in about the same number of moves.

2 The values of I for the patterns of Fig. 1 are .883 bits for patterns 1 & 4 ($1/5(2.0+0+1.+.415+1.0)$) and .556 bits for patterns 2 and 3 ($1/5(.415+0+1.0+.415+1.0)$). The value of E for this array is .724 bits ($1/5(.811+0+1.00+.811+1.00)$).

For the investigation of reinforcement effects we used answer sheets with a bimodal E_1 distribution;[3] i.e. the patterns were so constructed that each move had either a maximum expected informational outcome (white circle on half of the a patterns and black on the rest) or a low expected informational outcome (black circle on only one of the a patterns). These two classes of move will be called safe and gambling respectively. In order to establish differential reinforcement of S's choice among moves, the pattern arrays were so constructed that all gambling moves "paid off" on the same one pattern. Thus, relative frequency of reinforcement for gambling was readily manipulated through selection of the patterns used as problems. The results of several studies show both moves and strategy scores to be decreasing linear functions of the proportion of problems on which gambling is rewarded.

Comparison of group learning curves of strategy scores on successive trials obtained in an early study suggested an analogy to partial reinforcement in that groups given partial reinforcement for gambling seemed to be approaching the same asymptotic gambling strategy as groups given continuous reinforcement for gambling (problems in which gambling always paid off) but they did so at a lower rate. However, when all groups were given an additional series of problems in which gambling was never rewarded (extinction) there was very little evidence of differential resistance to extinction. Rather, all groups rapidly learned to adopt a safe strategy. Thus, manipulation of the proportion of trials on which gambling is rewarded does not seem to be equivalent to partial reinforcement. In view of the fact that an "unreinforced" gambling move does yield some information, and that choice of a safe move when a gamble would have been reinforced slows the ultimate solution by very little, the more appropriate analogy would seem to be to amount of reinforcement.

Very rapid transfer has also been found for groups which differed on initial training with respect to the number of patterns presented in the solution array. When groups trained on problems with 8, 16, or 32 patterns on an answer sheet were switched to problems with 32 possible answers there were no differences among groups either in moves or strategy scores. In fact, the only variable which we have found to have any differential effect upon transfer is the form of the E_1 distribution on the pre-transfer series. Unfortunately, we do not have much systematic data as yet on the nature of the effect of E_1 frequency distribution upon transfer.

It would seem intuitively obvious that when information is expensive S will try to make do with less. Moreover, it would seem reasonable that the type of information gathering strategy adopted should be influenced by S's past history of reinforcement: e.g. high information costs will increase the tendency to gamble if gambling

[3] Frequency distributions of E_1 such that for m/2 moves $E_1 = 1$ and for m/2 $E_1 = .544$.

has paid off in the past and vice versa. Although our data show trends in the expected direction neither information cost as a main effect nor the information cost x reinforcement history interaction is statistically significant. Score points in a competitive game were used in the above experiment. It may be that more meaningful (e.g. monetary) information costs would have a greater effect.

Summary. An experimental program designed to study diagnostic problem solving in relation to the problem itself is described. Three classes of independent variables were employed: informational aspects of the problem, training conditions, and information costs. The experimental task required S to determine which one of a number of patterns in a solution array was concealed beneath his problem board by obtaining information concerning as few of the elements as possible. Each S did a series of from 4 to 16 successive problems depending upon the conditions for his group.

Two response measures were considered in relation to the experimental variables. The first, number of information getting moves, is a more superficial measure of solution behavior which is determined by the more immediate specific informational aspects of a problem such as number of alternative solution patterns and specific pattern chosen as a problem. The quality of S's information gathering behavior is described by his strategy score. This measure tends to be relatively independent of specific informational aspects of the problem and to be influenced by more general informational characteristics such as the form of the distribution of information available.

With respect to learning variables such as number of trials and schedule of reinforcement both response measures are similarly affected. Learning, as inferred from decreased moves and more efficient strategies, takes place over a series of trials. Distinct strategies can be developed through a schedule of differential reinforcement. Transfer is positive and rapid. Finally, with respect to information costs, the present results are inconclusive.

REFERENCES

GLASER, R., DAMRIN, D. E. and GARDNER, F. M. The Tab Item: A technique for the measurement of proficiency in diagnostic problem solving tasks. *Educ. and Psych. Meas.*, 1954, 14, 283-293.

GLASER, R. and SCHWARZ, P. A. Scoring problem solving test items by measuring information. *Educ. and Psych. Meas.*, 1954, 14, 665-670.

GOLDBECK, R. A., BERNSTEIN, B. B., HILLIX, W. A., and MARX, M. Application of the half-split technique to problem solving tasks. *J. exp. Psychol.*, 1957, 55, 330-338.

JOHN, E. R. and MILLER, J. G. The acquisition and application of information in the problem solving process: An electronically operated logic test. *Behav. Sci.*, 1957, 2, 291-301.

MOORE, O. K., and ANDERSON, S. B. Search behavior in individual and group problem solving. *Amer. Sociol. Rev.* 1954, 19, 702-714.

BRIGHTNESS ENHANCEMENT WITH MICROSECOND PULSES[1]

W. L. GULICK

University of Delaware

Psycho-physical studies concerning visual perception are normally conducted under conditions wherein stimulation of the eye extends for indefinite periods of time. The use of intermittent stimulation in which the display of the stimulus is controlled temporarily has been restricted primarily to inquiries of fusion. There are, however, several visual phenomena which depend upon intermittent stimulation besides fusion (c.f.f.). Bartley (1936), for example, lists marginal flicker, course flicker, border flutter, glitter, and phase overlay as discrete perceptual experiences produced by undulatory stimulation below the rate required for fusion.

Perhaps the best known sub-fusion phenomenon in visual perception is the *Brücke effect*.[2] This effect refers to an enhancement of brightness at certain sub-fusion flicker rates which exceeds the brightness which would normally be perceived by an observer if the light source were steady.

Whenever a steady source of light is interrupted in such a way that the light and dark phases of the cycle are equal in duration, two perceptual changes occur. First, the magnitude of the sensation of brightness is reduced by one-half provided that the rate of interruption is higher than the rate required for c.f.f. (Talbot-Plateau law); and second, progressive reductions in flicker rate are accompanied by an enhancement of brightness which appears maximal for rates between eight and ten flashes per second (Bartley, 1939).

In his earlier work, Bartley demonstrated that the enhancement of brightness decreased as the light-dark ratio changed so that light occupied a larger proportion of the total light-dark cycle. Nevertheless, under conditions of interrupted transmitted light, "the point at which maximum brightness is achieved occurs at a constant flash rate regardless of the light-dark ratio" (Bartley, 1941, p. 137).

Recently Bartley, Paczewitz, and Valsi (1957) have shown that under certain conditions of intermittent photic stimulation there is a

[2] The term *Brücke effect* is the common denotation for the perceptual effect of brightness enhancement produced by flicker. To distinguish between reflected and transmitted light in producing this perceptual experience, Bartley refers to the brightness enhancement produced by the latter means as the *Bartley effect*. The more common usage is held to in this paper.

range of increased brightness which exceeds the level predicted from an extrapolation of Talbot's law, but falls short of the level one would expect under conditions of continuous stimulation.[3] When brightness from intermittent stimulation falls between these limits, Bartley et al. refer to this as the range of intermediate effectiveness (IR). The enhancement is too small to meet the conditions of the *Brücke effect*, and yet some enhancement does occur.

Based upon the early work of Bartley one might make two predictions: first, as the fraction of the total light-dark cycle devoted to light becomes smaller and smaller, the enhancement of brightness at subfusion flicker rates will become greater and greater; and second, the rate of flicker which produces maximum brightness enhancement is independent of the light-dark ratio. It should be pointed out, however, that both of these predictions are based upon the result of studies in which the light-dark ratios varied over a very narrow range (from 1:1 to 8:1).

The purpose of the present experiment was to consider the efficacy of these two predictions under conditions wherein the light-dark ratio was widely different from those employed previously. In particular, two questions were asked. First, does the *Brücke effect* occur when the light phase is exceedingly small as compared with the duration of the dark phase. Second, if the *Brücke effect* does not occur under these conditions, is there any enhancement of brightness over the level predicted by an extrapolation of the Talbot level which could be interpreted within the framework of Bartley's *intermediate range* (IR). In this instance, of course, the brightness would have a magnitude in excess of the extrapolated Talbot level but less than the brightness magnitude produced by the same intensity of light applied continuously.

METHOD

Subjects. Seven Ss, aged 18-21 years, were employed in this study. Each S met a visual acuity criterion of 1.0 (corrected), and each was naive about brightness enhancement and the purpose of the experiment.

Apparatus. The experiment was conducted in a darkened light-proof room. The visual field consisted entirely of two circular patches of light. The diameters of the light sources each subtended 1° of visual angle, and their centers were separated by a visual angle of 15°.

The light to S's right was produced by a photostimulator (Grass, Model PS-2) placed behind a diffusion screen. The duration of the light pulse equalled 10 microseconds, and the frequency of presentation was continuously variable from 1 to 100 flashes per second. It should

---

[3] Some writers, notably Bartley, refer to a Talbot level below c.f.f. Inasmuch as Talbot's law does not apply under conditions of flicker, the present author indicates that the level of brightness used as the reference for measures of brightness enhancement is actually an extrapolation of Talbot's law.

be pointed out that although the peak intensity of the photostimulator was equal to 93,750 candles, at a flash rate of 28 per second (approximate rate for c.f.f.) the Talbot level equalled 26 candles. Since the pulse duration of 10 microseconds was constant, variations in flash rate produced variations in light-dark ratios, and therefore, in Talbot levels as well.

The light to S's left was produced by an incandescent source placed behind a diffusion screen. Tinted filters were placed between this source and its diffusion screen in order to match the spectrum of the photostimulator source. The spectrum was practically flat from 400 to 700 millimicra.

The S controlled the intensity of the comparison light (the steady light to his left) with an autotransformer. In the first phase of the experiment S also controlled the frequency of the standard light (the intermittent light to his right).

Procedure. After acuity measures were completed, the S was seated at a table and his head was braced in a head-rest. The S viewed the stimuli monocularly with the preferred eye. Seven minutes were taken for dark adaptation during which the S was instructed as to how to vary the flash rate of the standard stimulus and the intensity of the comparison light. The S was allowed to look back and forth from one light to the other as many times as he wished.

In the first phase of the experiment the S set the intermittent light at fusion. Two ascending and two descending trials were used. Once the c.f.f. was determined, the E set the flash rate to equal the mean flash rate of S's four judgments. The S then adjusted the comparison stimulus to match the brightness of the fused light. Autotransformer dial settings were later calibrated for their concomitant brightness using a Macbeth illuminometer.

In the second phase of the experiment S was required to match with the comparison stimulus the brightness of the standard when the flash rate of the standard was set by E to equal 2, 4, 8, 16, 32, 64, and 100 flashes per second. The order of flash rates was varied systematically for each S, and for each flash rate each S made two brightness judgments.

To reduce kinesthetic cues, S was instructed after each setting of the autotransformer to remove his hand from the control knob and place it in his lap. Thereafter the E set the initial brightness of the comparison stimulus so that S would ascend or descend to reach equality in each flash rate condition. The S was allowed to reverse direction if he so wished.

Most Ss were able to make a brightness judgment in about 30 seconds. Two minutes elapsed between successive settings except on the occasion of the beginning of replication, when S rested for five minutes.

RESULTS

Data from this experiment show that some form of brightness enhancement occurred when the light-dark ratio approximated 1:5 x 10⁴. Summary data are presented in Figure 1 showing the mean brightness adjustment in apparent foot-candles for all Ss for each of the seven flash rate conditions (solid circles). Here it may be seen that the flash rates above c.f.f. (32, 64, and 100 flashes per second) produced brightnesses which were related linearly with the intensity of the intermittent stimulus as determined by the light-dark ratio according to the Talbot law. The second function (open circles) represents the Talbot level, and it represents the change in brightness for intensity values of the intermittent stimulus computed according to Talbot's law for the light-dark ratios which obtained under each of the flash rate conditions.
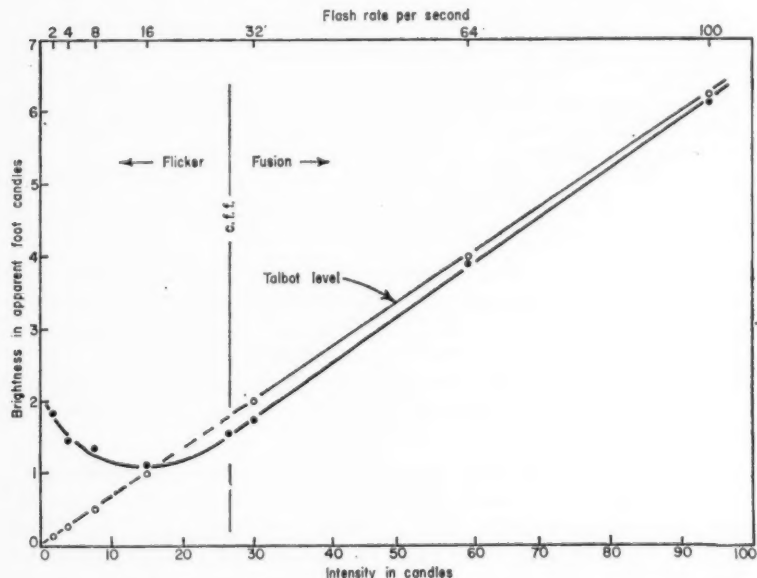
Figure. 1. The solid-circle function shows the mean brightness in apparent foot-candles for seven Ss of an intermittent light at each of seven flash rates. The open-circle function represents the Talbot level and indicates the intensity of the intermittent light as it was controlled by the light-dark ratios obtaining under each flash rate condition. The data obtained under flash rates above c.f.f. show that brightness is a linear function of light energy. Below fusion, using an extrapolated Talbot level (dashed line), it may be seen that brightness is not related linearly, or even directly, with light energy.

The brightness values obtained in this experiment confirm Talbot's law. For flash rates above c.f.f., brightness is a linear function of the intensity of the intermittent stimulus. The confirmation is reflected in the close approximation of the observations to the Talbot level. Under the conditions of slower flash rates (2, 4, 8, 16 flashes per second) the

perceived brightness deviated from the extrapolated Talbot level, thereby indicating that *below c.f.f. brightness is not related linearly, or even directly, to the total light energy striking the eye per second.*

An analysis of variance based upon brightness values indicated that the flash rate conditions resulted in significantly different brightness matches (.05 level of confidence). Moreover, since the *subject* by *condition* interaction was not a significant source of variance, it may be concluded that the enhancement of brightness at slow flash rates was typical of the Ss used in this experiment.

DISCUSSION

The relationship obtained in the present experiment between intensity and brightness is complex. For rates of interruption above c.f.f. brightness appears to be related linearly with intensity, as controlled by variations in the light-dark ratio. These data simply confirm Talbot's law. For rates below fusion, however, the relationship between the amount of light energy and brightness is very different. In reference to Figure 1 it may be noted that 2 flashes per second (20 microseconds of light) and 32 flashes per second (320 microseconds of light) produced equivalent brightnesses. Moreover, for flash rates between 2 and 16 per second the relationship of light energy to brightness becomes inverse.

If the brightness below fusion is compared with the brightness predicted by an extrapolation of Talbot levels, then the data from the present experiment may be considered as demonstrating some form of brightness enhancement since the perceived brightness always exceeded the predicted value. In no instance, however, did the brightness exceed that which would have obtained had the source been steady rather than intermittent. Accordingly, this is *not* the *Brücke effect.* Furthermore, enhancement was maximal at the slowest rate studied (2 flashes per second) rather than at rates between 8 and 16 flashes per second where the *Brücke effect* is reported to be at its maximum.

The question may now be raised as to whether or not this deviation can legitimately be classified as brightness enhancement. When the rate of presentation of the flash of light was as slow as two per second, the subject seemed able to judge each flash independently. The brightness of a single flash is partly dependent upon the state of the eye at the time of presentation. At slow rates the processes involved in dark-adaptation would have relatively long periods in which to adapt the eye as compared with the periods accompanying faster rates of presentation. Accordingly, the *off* period is especially important under conditions of intermittent stimulation. Apparently at sub-fusion rates the brightness of an intermittent stimulus is maximal under those con-

ditions which allow the most recovery of photochemical substances at the retina between periods of stimulation.

The brightness of a steady light viewed against a dark background is probably dependent upon different levels of neural excitation in the visual system produced by the stimulus object and its background. If this is so, then brightness depends primarily upon the existence of a *difference* in stimulation at discrete retinal areas as a result of an unequal distribution of light energy in the visual field. The physiological correlate of brightness under conditions of steady continuous stimulation is therefore a *spatial* one. However, under conditions of intermittent stimulation a *temporal* correlate is added. The brightness perceived by stimulation of a particular retinal locus depends now not only upon differences in levels of excitation with other retinal loci, but upon differences in the levels of excitation within the area in question as it fluctuates through time.

When the rate of interruption is fast, the temporal fluctuations in excitation are minimized because of the rapidity with which the light and dark adaptation processes alternate (Hecht, 1934). Accordingly, the physiological cue to brightness would involve differences in levels of excitation as they are distributed spatially over the visual system. At very slow rates of presentation it appears as though the visual system is quite sensitive, probably because of the favorable state of the eye in regard to dark-adaptation. The added cue of a temporal fluctuation within the area under stimulation may account for the apparent increase in the efficiency of brightness perception during slow rates of stimulus interruption.

In addition to a consideration of light- and dark-adaptation at the retinal level, it is probably necessary to inquire of the nature of the response of various fiber types found in the mammalian eye in order to determine if the phenomenon observed in this experiment might be partly explained at post-receptor levels.

Granit and Therman (1934, 1935) and Hartline (1938) noted that those fibers which normally respond to the cessation of a light stimulus can be inhibited by re-illumination. The temporal sequences noted suggest that the flash rates producing minimum brightness in the present experiment are in the same range as those producing maximum inhibition of the *off* type fibers. If one assumes that the *off* discharge is important to brightness perception, then the observed inverse relationship obtained in this experiment may have been due to an inhibitory effect of neural origin.

Finally, mention should be made of Bartley's alternation of response hypothesis (1938, 1942, 1952). According to his interpretation, any factor which would work toward distributing neural activity throughout the light-dark cycle ought to reduce brightness enhance-

ment. When the light occupies an exceedingly brief portion of the total cycle, as in the present experiment, enhancement should be greatest. This is indeed the case on the basis of the present findings. The major difficulty, however, is that Bartley et al (1957) report data showing that maximum enhancement occurred when the light occupied a fraction of the total cycle equal to 0.3, and that smaller fractions of light produced *less* rather than more enhancement. The data from Bartley et al. seem to be inconsistent both with the findings of the present experiment and with the alternation of response hypothesis.

SUMMARY AND CONCLUSIONS

Each of seven Ss matched with a steady light the brightness of an intermittent light under seven flash rate conditions. At flash rates above c.f.f. the brightness of the intermittent light was linearly related to the light energy striking the eye per second, as determined by Talbot's law. Below fusion brightness deviated from the predicted level as determined by an extrapolation of Talbot levels. Brightness was inversely related to light energy striking the eye per second.

This phenomenon could be interpreted as brightness enhancement if the extrapolated Talbot levels are used as a reference. The light-dark adaptation cycle and inhibition of the *off* discharge were offered as possible explanations of this phenomenon.

Based upon this experiment it is concluded that enhancement of brightness can occur even under conditions wherein light-dark ratios approximated $1 : 5 \times 10^4$, but the *Brücke effect* cannot be demonstrated under these conditions. Moreover, the rate at which maximum brightness is achieved does not occur at a constant flash rate which operates independently of the light-dark ratio.

REFERENCES

BARTLEY, S. HOWARD. The basis of the flicker in the visual field surrounding the test-object. *J. exper. Psychol.,* 1936, 19, 342-350.

BARTLEY, S. HOWARD. A central mechanism in brightness discrimination. *Proc. Soc. Exper. Biol. and Med.,* 1938, 38, 535-536.

BARTLEY, S. HOWARD. Some factors in brightness discrimination. *Psychol. Rev.,* 1939, 46, 337-358.

BARTLEY, S. HOWARD. *Vision.* New York: Van Nostrand, 1941.

BARTLEY, S. HOWARD. Visual sensation and its dependence on the neurophysiology of the optic pathway. *Biol. Symposia,* 1942, 7, 87-106.

BARTLEY, S. HOWARD. Vusual response to intermittent stimulation. *Opt. J. Rev.,* 1952, 89, 31-33.

BARTLEY, S. HOWARD, PACZEWITZ, G., and VALSI, E. Brightness enhancement and the stimulus cycle. *J. Psychol.,* 1957, 43, 187-192.

GRANIT, R., and THERMAN, P. O. Inhibition of the off-effect in the optic nerve and its relation to the equivalent phase of the retinal response. *J. Physiol.*, 1934, 81, 47 p.

GRANIT, R., and THERMAN, P. O. Excitation and inhibition in the retina and in the optic nerve. *J. Physiol.*, 1935, 83, 359-381.

HARTLINE, H. K. The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. *Amer. J. Physiol.* 1938, 121, 400-415.

HECHT, S. Vision: II. The nature of the photoreceptor process. In C. Murchison (Ed.), *Handbook of experimental psychology.* Worchester, Mass.: Clark Univer. Press, 1934. Pp. 719-751.

THE RELATION OF EXTRANEOUS VISUAL
STIMULI TO APPARENT SIZE[1]

ALVIN G. GOLDSTEIN

University of Missouri

The data of a visual perception experiment[2] were interpreted to show that, in a display crowded with items similar in form but of only two sizes, it is extremely difficult to "see" the two groups as independent. A perceptual interaction among the stimuli was thought to be responsible for this result. It was hypothesized that visual size judgments are affected by the number and position of adjacent stimuli. The present report describes the results of several experiments which were designed to test this general hypothesis.

METHOD

All experiments employed the psychophysical method of constant stimuli with two categories of judgment. Two comparison stimuli, a variable (V) and a fixed standard (S), were simultaneously presented to observers (Os) and they were required on every trial to select the larger of the two. Notice that "equal" or "doubtful" responses were not permitted. Extraneous (E) stimuli were sometimes presented in the vicinity of V but the O's were informed that the E-stimuli were not involved in the judging task. The interaction of E-stimuli with S- and V-stimuli was the main concern of the studies.

The experiments were designed to determine whether apparent size of the comparison stimuli was a function of: (a) the type of figure, that is, when S-, V-, and E-stimuli are all the same shape; (b) the number of E-stimuli; (c) the patterns or positions of the E-stimuli; (d) the distance between E- and V-stimuli; and (e) the type of E-stimuli, that is, when S- and V-stimuli differ in shape from the E-stimuli. For all experiments reported here the general procedure, stimulus presentation, etc., was identical. The comparison stimuli each subtended approximately 0.75° visual angle and both were at eye level separated by 2.5° visual angle. S and V were either presented alone, or with one to five E-stimuli situated in the 10 spatial positions shown in
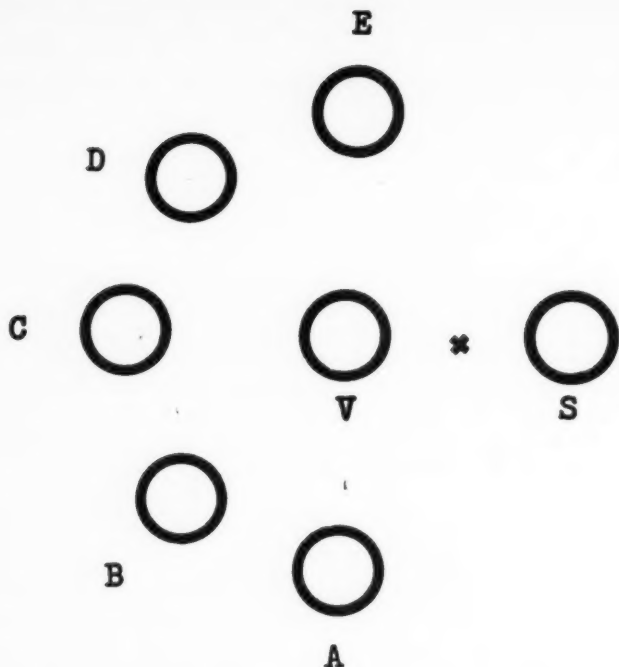
Fig. 1. Spatial positions of E-stimuli and comparison stimuli, S and V, as seen by the O. The circles in the figure are replicas of those used in one experiment; corresponding positions were occupied by squares and lines in the other experiments. In one-half the trials, the display was reversed from left to right.

Fig. 1. and its mirror image.[3] With only one exception (e), E-stimuli were always identical to V in *all physical aspects* and were situated approximately 2.5° or 5° visual angle from V. Three classes of stimuli were used: circles (annular rings), line lengths, and outline squares. Stimuli were white figures (0.5 ft-Lamberts) on a dark (0.02 ft-Lamberts) ground. With S physically equal to V (S=V), 32 slides were developed, using all combinations of zero to five E-stimuli (e.g., one slide with S and V alone, 5 slides with a single E-stimulus, etc.). Similarly, two sets of 32 additional slides could be made with S approximately 1 per cent larger (S>V) and smaller than V (S<V). In certain experiments all 96 slides were employed whereas in other experiments only selected ones were used.

Stimuli were presented for 7 to 8 seconds in a semi-dark room by a 3¼ x 4 inch projector (American Optical Co.) using a 300 watt bulb

[3] The small "x" shown in Fig. 1 between S and V served to help the Os quickly locate the comparison stimuli. Data collected without the "x" clearly indicate that it has no effect on judgments.

(10 inch, f4.5 lens with 17 *mm* diaphragm) to groups of 6 to 12 *O*s. Responses were written on specially prepared answer sheets. Normal unrestricted binocular viewing was employed at all times. Although the total number of trials varied among the experiments, each *O* judged every slide at least twice, once with E-stimuli in the right visual field and once in the left. The number of trials where S>V always equalled the number where S<V. And, as previously noted, when S=V *O* still had to select the larger of the two stimuli. Thus, in the long run, S and V each will be selected at the larger stimulus 50 per cent of the time, unless there is some other factor causing a bias for either S or V.

RESULTS AND DISCUSSION

Figure 2 (bottom) presents the mean per cent of S>V judgments as a function of the number of extraneous stimuli in three experiments utilizing different classes of stimulus figures. The means are based on all the judgments of 42 different *O*s in each experiment. Without E-stimuli S>V judgments are close to 50 per cent. With one E-stimulus there is a significant (minimum $p<0.05$) increase in the per cent S>V judgments.[4] Adding more than one E-stimulus, however, has no further important effect on judgments. Plotting results for the S=V condition alone produces a similar set of curves (Fig. 2, top). These data confirm the findings of an earlier study (Goldstein, 1959) which employed circle stimuli and suggested that 6 E-stimuli were as effective as 13 in eliciting the size reduction effect. Although it is tempting to make the categorical statement that the influence of one is equal to the influence of several E-stimuli, it would not be a proper interpretation of the data. In order to obtain that kind of information, it would be necessary to perform a quantitative study of the *amount* of E-stimuli influence as a function of the number of E-stimuli.

The data were also analyzed with regard to the spatial position of the E-stimuli, and it is safe to say that position in the visual field is a contributing factor nothwithstanding the fact that visual fixation was not employed. With number of stimuli held constant, some displays yielded more than 90 per cent S>V judgments as compared to 48 per cent for other displays. Very briefly, the most effective displays in the three studies involved the following positions; A, E, D, and B in the left visual field and E in the right (Fig. 1).

Another experiment (36 *O*s, 2,160 trials) was performed employing the circles as stimuli to determine the role of retinal distance between E- and V-stimuli. E-stimuli were situated approximately 2.5° and 5° visual angle from the V-stimulus. With S=V, the mean S>V judgments for the 2.5° conditions was 73 per cent and for the 5° conditions, 65 per cent, a difference which was not significant at the 0.05 level.

[4] Although the graphed data represent responses and not individual *O*s, chi square analyses of each *O*'s response strongly support these findings. S was judged larger than V in significantly more than 50 per cent of the responses made by 93 of the 126 *O*s. Of these 93 chi squares, 67 were significant beyond the 0.001 level, 16 beyond the 0.01 and 10 beyond the 0.05 level.
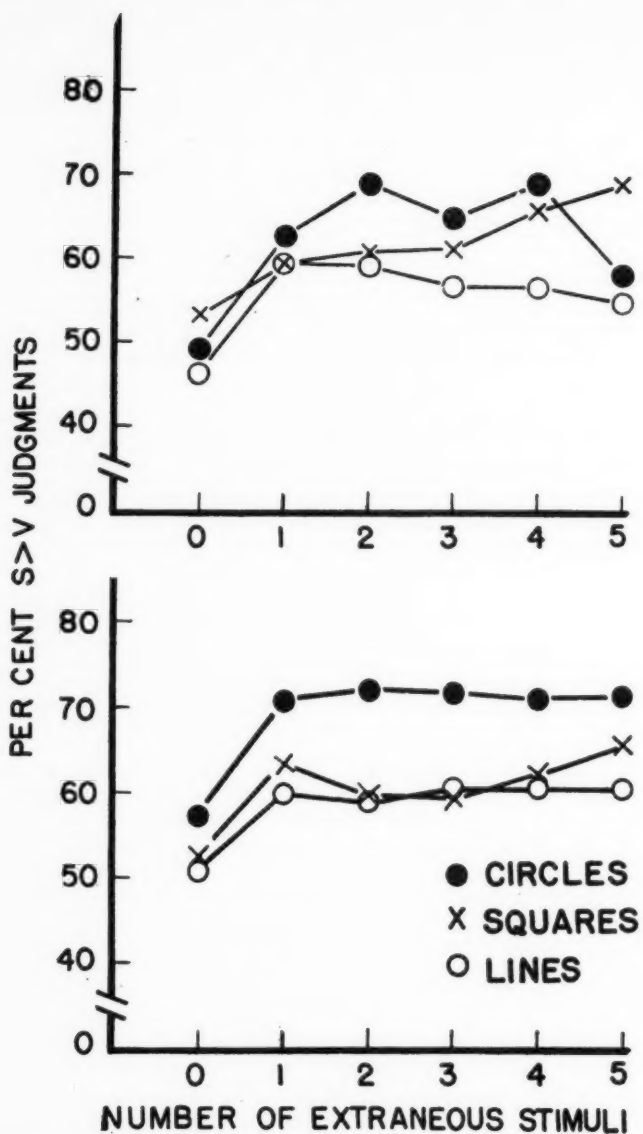
Fig. 2. Size comparison judgments with three types of stimulus shapes as a function of the number of extraneous stimuli. (Bottom) Mean per cent S greater than V judgments when S was actually larger, equal, and smaller than V. (Top) Mean per cent S greater than V judgments when S was actually equal to V.

For the trials with one E-stimulus, however, increasing the V-E distance resulted in a large and significant difference (70 vs. 54 per cent) which, along with other data in the experiment, suggests that distance is a less important factor as the number of E-stimuli increase.

In problems of this type, "contrast" is often offered as an "explanation" of the results. In the present instance contrast could be invoked by assuming that the E-stimuli appear larger than they actually are due to their peripheral location on the retina, and therefore, by contrast the V-stimulus appears smaller than it actually is. Neglecting all the other difficulties which this hypothesis encounters, the evidence points to a rather limited role for either the size or the type of the E-stimulus. This evidence was obtained in an experiment (14 Os, 1,050 responses) where comparison stimuli were identical-sized outline circles and the E-stimuli differed from V in kind (outline-squares, lines, discs), and sometimes in size of one dimension (area, length or diameter). Without E-stimuli, S>V judgments came to 48 per cent of the responses. With outline-squares as E-stimuli S>V judgments increased to 76, 77, and 74 per cent when E was equal to, larger and smaller in area than V, respectively. The discs, which gave off much more light than any other E- stimulus, resulted in 80 and 74 per cent of S>V judgments when they were equal to and smaller in diameter, respectively, than the V stimulus. Two sizes of outline circles, both smaller in diameter than V, resulted in 71 and 76 per cent S>V judgments. Two sets of line lengths much smaller in length than the diameter of the V stimulus each resulted in 59 per cent S>V judgments (not significantly different than 50 per cent) even though the lengths differed from the diameter of V by unequal amounts.

Two perceptual theories by Obonai and Köhler may appropriately be considered in relation to the results reported here. Obonai has proposed the concept of psycho-physiological induction where ". . . the areas surrounding the stimulated part of the retina undergo a change in excitability when a light stimulates one part of the retina . . . Such an excitability appears as various sensory or perceptual properties such as sensitivity, brightness, color, spatial extent, propensity, etc." (Obonai, 1957, p. 3). In order to handle overestimation and underestimation of perceptual extent, and presumably other types of "errors" of judgment, Obonai introduced the assumption that overestimation is due to a direct excitatory process whereas underestimation is a result of an opposing process (Obonai, 1957, p. 8). Although psycho-physiological induction has been called a theoretical explanation, it is much too restricted to be designated as anything more than an hypothesis. Because it is so limited, and in addition, so vague, the only inference that can readily be made is that the presence or absence of a non-focal visual stimulus will have an effect on the perception of a focal stimulus. (This relationship will obtain when both the inducing and the focal stimuli are presented either simultaneously or separated

by a short time interval.) The direction of this effect, i..e, whether the "perceptual error" is larger or smaller than the "true" size of the stimulus is not deducible from the theory. There is the double implication, however, that overestimation and underestimation are functions of particular aspects of the inducing stimulus and that these two phenomena can always be obtained if the inducing stimulus is varied. These concepts can be applied to the results reported here, but it is doubtful whether *at this stage* there is any advantage to be gained. Neither the direction of the effect nor the relationship between the number of extraneous (or inducing) stimuli and the intensity of the phenomenon could have been predicted. The implication that there should be a shift from overestimation to underestimation with a change in the inducing stimulus does not appear to be supported by the data since the V-stimulus in all experiments was always judged smaller than the S-stimulus even though the E-stimuli were varied along several dimensions. It must be remembered, however, that there was no systematic attempt to make a fair test of this theoretical point. In addition, although the theory aims at explaining perceptual response on a physiological level, it offers little help in understanding the mechanism of the present phenomenon nor does it suggest to the author the direction in which research should proceed.

Köhler and Wallach's "brain field" theory and its phenomenological expression, the figural aftereffect (1944) is well known and will not be described here. It also is of doubtful usefulness with regard to the present findings because of a crucial point. Greatly simplified, the theory states that when two stimuli are serially presented to the same (or nearly the same) cortical area, the first stimulus creates electrical and physiological conditions within this area which affects the perception of the second stimulus. According to the theory, judgments of size are affected by a particular set of conditions; the position on the cortex of the second stimulus must coincide with that of the first, otherwise an apparent shift in position will occur, instead of a change in apparent size. In the experiments reported here there was no possibility that *only* the V-stimulus was falling on a "saturated" part of the cortex. A prior stimulus could affect a subsequent one only from one trial to the next, but the S- *and* the V-stimulus were always presented together and therefore any prior cortical effects must be assumed to influence both stimuli. Furthermore, visual fixation was not required of the Os, thus making it highly unlikely that the stimuli fell on the same cortical area from trial to trial. There is almost no doubt that the procedures employed in the experiments did not permit the occurrence of "figural aftereffects" consequently, it must be concluded that the results are not interpretable in terms of a satiation theory.

SUMMARY AND CONCLUSIONS

The apparent size of one of two simultaneously presented comparison figures was reduced when viewed with one to five extraneous

visual stimuli. These extraneous stimuli in one set of experiments were identical in size and shape to one of the comparison figures. Similar results were obtained when the extraneous stimuli differed in kind and size from the comparison figures and also it may be concluded that, within the limits discussed, the size reduction effect (a) is a general phenomenon which can be demonstrated with several types of visual shapes; (b) elicits almost the same per cent S>V judgments with one or several E-stimuli; (c) is affected differentially by the relative spatial positions of the E-stimuli; (d) is dependent to some degree upon the size of the visual angle between V- and E-stimuli; it is suggested that there is an interaction between the number of E-stimuli and their distance from V; (e) occurs when the E-stimuli are different in kind and size from the V-stimulus, but the amount of this influence probably depends on the particular class of stimulus.

REFERENCES

GOLDSTEIN, A. G. Size comparison judgments as affected by extraneous visual stimuli. Paper presented at 1959 meeting of Midwestern Psychology Association.

KÖHLER, W., and WALLACH, H. Figural after-effects: an investigation of visual processes. *Proc. Amer. Philos. Soc.* 1944, 88, 269-357.

OBONAI, T. The concept of psycho-physiological induction. *Psychologia*, 1957, 1, 3-9.
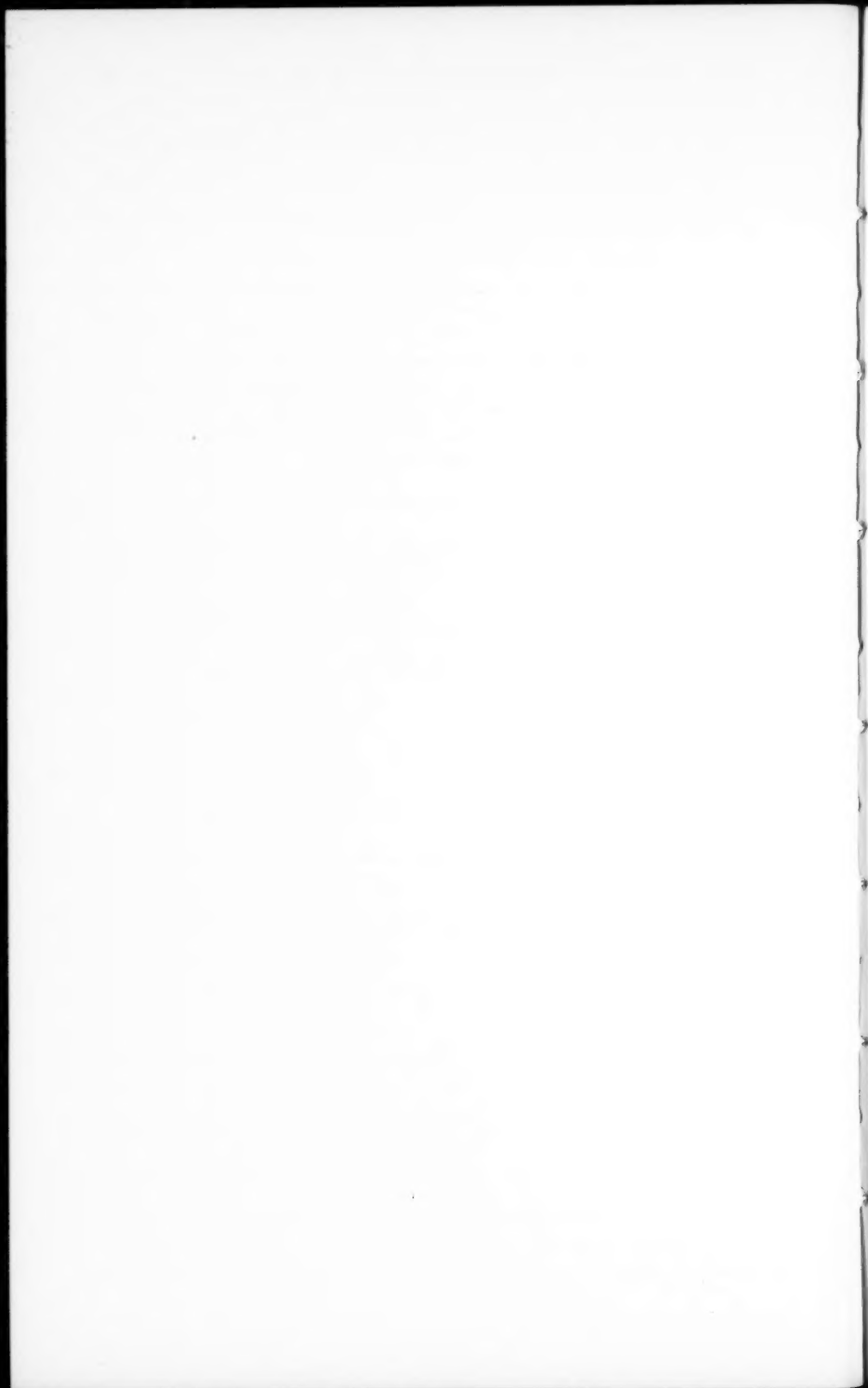
REACTION TIME TO ONSET AND CESSATION OF A VISUAL STIMULUS[1]

JACK D. RAINS
University of Arizona

A topic of some interest to experimental psychologists in the early decades of this century was that of the relationship between reactions to onset and to cessation of lights. Teichner (1954) in his review of studies of simple reaction time concludes that the matter "seems far from settled."

One of the earliest experiments on the topic was conducted by Woodrow (1915). Employing three different intensities and seven subjects, five of them practiced, Woodrow concluded that no difference existed between the two types of reaction. Jenkins (1926) using unpracticed subjects found shorter reactions to cessation than to onset of visual stimulation. Both of these experiments provided no fixation point for their subjects, a factor which, it might be argued, would mitigate for longer reactions to onset of stimulation since subjects might consume some time in directing their attention to the light when it was illuminated. It was this consideration (together with certain others) which led Steinman (1944) to study reactions to increments and decrements in intensity of visual stimuli. On the whole, she found shorter reactions to decrements in intensity at all levels. Although she did not examine the limiting case where the change is from no prevailing illumination to a high level and from a high level of prevailing illumination to darkness, her graphs appear to show the approach of the increment and decrement curves at the higher levels of absolute change and she suggests that at these higher levels the curves might eventually join.

In the experiment which follows, we have investigated this limiting case—that of onset and cessation of a bright, large flash against a dark field—to determine if differences in reaction time to onset and cessation exist at this level of inquiry.

METHOD

Apparatus

The response timing apparatus was a custom designed unit which employs a Berkeley counter to time reactions to a tenth of a millisecond. The unit contains a provision for delivering a flash or extinguishing a light for controlled durations between .001 and 10 seconds.

The flash was delivered by a Sylvania Glow Modulator (R 1131C). This tube, which has a rise and decay time of less than .1 milliseconds,

was housed in one arm of an L-shaped visual discriminator, depicted in Fig. 1. The light from the crater of the tube passed first through a collimating lens, then through a neutral density filter and a double-prism
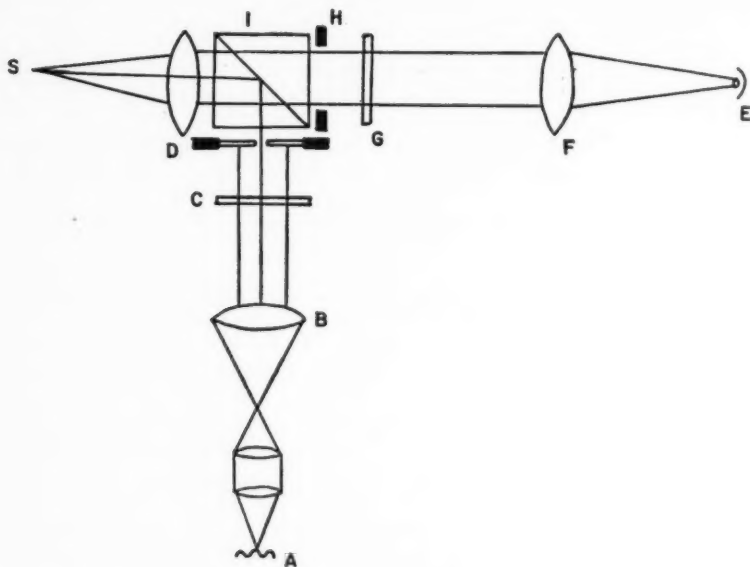

Fig. 1. Schematic drawing of visual discriminator. A—light source for fixation point; B—collimating lens; C—red filter; D—pinpoint field stop; E—stimulus glow tube; F—collimating lens; G—neutral density filter; H—field stop for stimulus light; I—double-prism beam splitter; S—subject's eye.

beam splitter, and finally through a lens which focussed the light in such a way as to provide the subject with a Maxwellian view. The other arm of the discriminator was used to provide a red fixation point which remained illuminated in the center of the field throughout all trials. The subject viewed the field through an artificial pupil. The intensity of the light was calibrated with a Macbeth illuminometer at 344 millilamberts. It subtended a visual angle of 12 degrees, 44 minutes.

Subjects

Three well-trained subjects were used in the experiment. They had all received several thousand practice trials prior to the experiment.

Procedure

During the trials, S was housed in semi-darkness in a small portable room for about five minutes before he began reacting. Several practice trials were administered prior to each block of trials and S indicated when he felt he was sufficiently "warmed up" to begin giving data.

In the onset condition, S was able to see only the red fixation point

when he positioned his eye before the artificial pupil. The experimenter gave a ready signal, S indicated his readiness, and from one to three seconds later, the flash was delivered.

In the cessation condition, S was able to see both the stimulus light and the fixation point at the start of the trial. Ready signals and foreperiod conditions were identical to those in the onset condition and S reacted when the light was extinguished. Thus, S viewed the fully illuminated field for about two to four seconds prior to reacting.

Four blocks of 25 trials each were presented daily to each subject. The trials were presented in one sitting in an ABBA order and, on other days, as a counterbalance, trials were presented in a BAAB order.

A few data were discarded. Specifically, if the subject felt that he had anticipated the stimulus or had been unduly sluggish in responding, he communicated this feeling to the experimenter who recorded the reaction but did not enter it in the computations.

TABLE 1

REACTION TIMES IN MILLISECONDS TO ONSET AND CESSATION OF LIGHT BY SUBJECT AND DAY

	Subject JR				Subject HK				Subject NB			
	Cessation		Onset		Cessation		Onset		Cessation		Onset	
Day	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	165	14	164	12	150	12	150	8	173	12	181	27
2	168	12	162	13°	152	21	146	11	170	11	181	9
3	160	10	159	10	148	15	141	17°	173	13	169	9
4	162	9	167	11°	142	16	153	12°	166	9	167	9
5	157	10	157	12	148	17	139	20°	172	8	166	7°
6	159	10	164	12°	145	11	142	12	177	14	176	15
7	161	8	166	10°	141	10	139	11				
8					144	14	145	20				

°Significant at .05 level

RESULTS AND CONCLUSIONS

The results of the experiment are summarized in Table 1. Regarding differences between onset and cessation reactions at these parameters, it may be concluded that no difference exists. It may be added, however, that preliminary exploration suggests that differences may be apparent when a small, dim, peripheral flash is employed.

In thirteen of the 21 replications, no significant difference between the two types of reaction was found. In eight experiments, differences were significant at the .05 level. Of these eight rejections of the null hypothesis, four were in one direction and four in another.

On the average, 21 statistical tests performed at the .05 level of significance would yield one significant difference due to chance. A finding of eight significant differences is surprising, particularly when the eight are split equally in the two tails of the distribution. Examination of graphic day-by-day plots of the means of the two types of reaction demonstrated no systematic patterning which would allow us to hypothesize an underlying cause for these eight significant differences. Since these subjects were highly reliable (as evidenced by their low standard deviations), they were extremely sensitive indicators of slight fluctuations in their physiological and psychological states. These fluctuations are assumed to account for the observed differences.

SUMMARY

An experiment was conducted to determine if differences exist in reaction time to onset and cessation of a large, bright, foveal light. Out of 21 replications of the experiment, eight differences, split equally between the two types of reactions, were significant. These findings are attributed to the trained subjects' sensitivity to momentary fluctuations in their physiological and psychological states. It is concluded that no differences between the two types of reaction are apparent at these parameters.

REFERENCES

JENKINS, T. N. Facilitation and inhibition. *Arch. Psychol., N. Y.*, 1926, No. 86, 1-56.

STEINMAN, ALBERTA R. Reaction time to change compared with other psychophysical methods. *Arch. Psychol., N. Y.*, 1944, No. 292. 1-60.

TEICHNER, W. H. Recent studies of simple reaction time. *Psychol. Bull.*, 1954, 51 (2), 128-149.

WOODROW, H. Reactions to the cessation of stimuli and their nervous mechanism. *Psychol. Rev.*, 1915, 22, 423-452.

GENERALIZATION OF EXTINCTION ON THE SPECTRAL CONTINUUM[1]

WERNER K. HONIG

Denison University

Compared to gradients of acquisition, the generalization of extinction has received little attention. Nevertheless, it has been assumed by some theorists (e.g. Spence, 1937; Hull, 1953) that primary decremental gradients of extinction (or inhibition) not only enter into the determination of behavior, but are of the same form as gradients of acquisition (or excitation), although they are of course inverted. This assumption was based on a few early studies of these gradients with classical conditioning. Generalization of extinction was reported by Pavlov (1927), where the response was salivation and the dimension of generalization either the frequency of a tone or the location of a tactile stimulus on the back of a dog. Bass and Hull (1934) found that extinction of the GSR generalized decrementally along the backs of human beings. Hovland (1937) used the same response with tonal frequency serving as the continuum for the CS.

In the design typical of this kind of study, Ss are first reinforced equally on a series of values along a generalization continuum. Each CS is then tested briefly without reinforcement to determine post-acquisition response strength which ideally is equal for all the stimuli. One CS (S^e) is then presented for a number of trials without reinforcement to induce extinction. Finally, the original stimulus series is again presented in a generalization test without further reinforcement. The gradient of extinction is obtained by comparing response strength at each value on the first and second tests.

Unfortunately, this design, when used in the classical situation, involves some procedural difficulties. The pre-extinction test, however brief, produces some extinction that could affect the gradient, for the UCS must be omitted on such trials. In the studies with human beings, the extinction of the GSR was so rapid that retraining was presented for the stimulus series in alternation with re-extinction and retesting. This would in effect permit discrimination learning between the S^e and the other stimuli, and as Kling (1952) points out, would result in discrimination gradients rather than pure gradients of extinction.

Studies using instrumental behavior are very few in number. Kling (1952) used independent groups, training Ss in each group to respond positively to two white circles of different size, and then extinguishing with one circle. The difference in size between the circles was systematically varied between groups. It was found that the degree to which the extinction generalized to the non-extinguished circle depended on this size difference. In spite of the use of independent groups, Kling found that extinction to S^e rapidly led to an extinction of all running, and he was forced to alternate extinction trials on S^e with trials in which running was reinforced to a black square. In effect, extinction was obtained with a discrimination training, rather than a "straight extinction" procedure. Kling does not present any gradients of acquisition with which the extinction gradients can be compared.

The operant technique used in the present study avoided the difficulties discussed above, due particularly to the advantages of using partial reinforcement in the acquisition of the response. The pre-extinction test without reinforcement at all training values could be eliminated, since response rates could be obtained during training. Furthermore, extinction could be attenuated to such a degree that retraining, re-extinction and retesting periods were not necessary for the determination of consistent gradients. Finally, each animal was tested on each stimulus value, thus avoiding the problems associated with determining gradients from independent groups.

The general technique is based on the work of Guttman and Kalish (1956), which provided gradients of acquisition for pigeons on the same continuum and with the use of a very similar procedure. A comparison of acquisition and extinction gradients obtained under similar conditions is therefore afforded by the present study.

METHOD

Subjects. The Ss were 15 experimentally naive pigeons of various breeds, maintained at about 80% of their free-feeding weight.

Apparatus. An automatic key-pecking apparatus, similar to that of Guttman and Kalish, was used. The only source of illumination in S's box was the pigeon's key, aside from a magazine light that was briefly lit during reinforcements. The light source was a Bausch and Lomb monochromator with a diffraction grating of 1200 lines/cm. Small portions of the spectral dispersion could be selected for transmission through an "exit slit" by rotation of the grating. An image of the exit slit was directed at the pigeon's key. The width of the transmitted band was about 16.5 Mμ; this provided a uniform hue of high purity and saturation for the human observer. The stimulus values presented to S will be designated by the middle value of the spectral band falling on the key. No attempt was made to equate for brightness differences among the stimuli. A number of studies (e.g. Guttman and Kalish,

1956; Honig, Thomas, and Guttman, 1959; Hanson, 1959) indicate that
brightness differences are relatively unimportant in the determination
of spectral generalization gradients.

Procedure

Preliminary training. Day 1) S was allowed to eat for three min.
from the open food magazine. Day 2) S was trained to eat during the
magazine cycle. The magazine was closed slowly at first and then with
increasing speed until the regular five-sec. cycle was attained. Day 3)
S was given 25 "free" magazine cycles to insure reliable eating when-
ever food was presented. Day 4) S was taught by the method of suc-
cessive approximations to peck at a key illuminated by a 570 Mμ stim-
ulus. Immediately after a few operant responses had been obtained,
the light beam was interrupted and the stimulus was changed. S was
presented with 510, 540, 600, and 630 Mμ stimuli and permitted to
obtain a few continuous reinforcements for pecking at each of these
values. Day 5) S obtained five continuous reinforcements with each of
the 13 stimulus values used in VI training and presented in a random-
ized order.

Variable interval (VI) training. Thirteen stimuli ranging from 510
to 630 Mμ in 10 Mμ steps were presented to S during ten daily sessions
of VI training. Each session consisted of 39 one-min. periods of stim-
ulus presentation alternating with 10 sec. of blackout to permit chang-
ing of the stimulus and recording of responses. Three randomized blocks
of the 13 stimuli were presented during a session. In the course of the
ten sessions, each stimulus was therefore presented for 3 min. per ses-
sion, or a total of 30 min. The VI schedule had a mean inter-reinforce-
ment interval of about 50 sec. S could receive from zero to three rein-
forcements during a given minute. It was assumed that with a random-
ized order of stimulus presentations, responding would be reinforced
about equally for each spectral value in the course of ten sessions.

Extinction to 570 Mμ. On the day after the last training session
the 570 Mμ stimulus (S^e) was presented while reinforcements were
withheld. Six Ss in Group 1 received 40 consecutive minutes (with
blackout periods, as before) of extinction in one session. Six Ss in Group
2 received two such sessions, or a total of 80 min. of extinction. Three
Ss in a Control Group received no extinction but proceeded directly to
the generalization test.

Generalization testing. On the days following extinction, general-
ization testing was carried out in a manner similar to the VI training,
except that reinforcements were withheld, and the stimulus presenta-
tion periods were 30 rather than 60 sec., again alternating with 10-sec.
blackouts. Six randomized blocks of the 13 stimuli were given in each
session. Groups 1 and 2 were given two such sessions, with the inter-
vening day permitting some spontaneous recovery. The Control Group,
which had received no previous extinction, was tested for four days.

Reconditioning and Retesting (Re-extinction). All Ss in Groups 1 and 2 were reconditioned in two sessions identical to the testing sessions except that reinforcements were introduced on the VI schedule used in training. The shorter periods were used so that a closer study could be made of changes in the gradient during the course of reconditioning. Retesting was thereupon carried out under extinction for two sessions in a manner identical to the first generalization test. No specific extinction intervened between retraining and retesting.

RESULTS

The data are presented in Figs. 1, 2, and 3 as mean number of responses per minute for Group 1, Group 2, and the Control Group respectively. The mean rates for each of the stimulus values for the last four sessions of VI training (at which time the overall response rate had
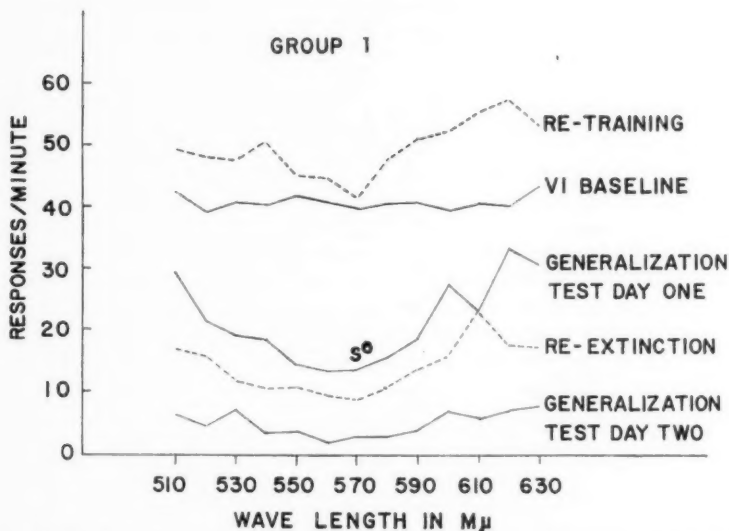
Fig. 1. Response rates for Group 1 for the end of VI training, for generalization testing, retraining, and retesting (re-extinction).

leveled off) provide a baseline for comparison with the generalization gradients. It is clear that the VI rates are very similar for all stimulus values in each group. The Ss did not "prefer" any particular value or range of values in each group, even though these ranged from bluish-green (510 Mμ) to red (630 Mμ).

The mean rates for the two days of initial generalization testing are separately presented for each day for Groups 1 and 2. There is an orderly gradient with its minimum at or near 570 Mμ for both groups.
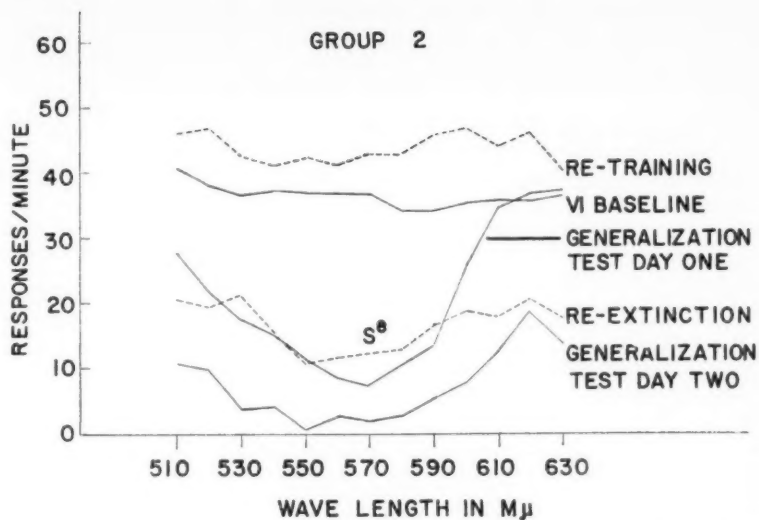
Fig. 2. Response rates for Group 2 for the end of VI training, for generalization testing, retraining, and retesting (re-extinction).
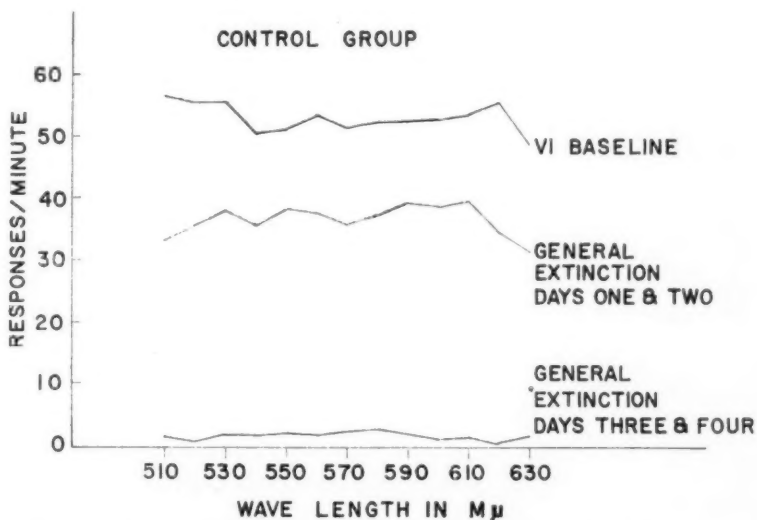

Fig. 3. Response rates for the Control Group for the end of VI training, and for general extinction on all stimulus values.

The gradients obtained on the second day are lower than those from the first, showing the cumulative effects of extinction incurred in the

course of testing. The gradients from the first day also are subject to this effect; the initial rates obtained during the first three blocks of test stimuli on the first half of Day 1, though not shown on the figures, reached the VI baseline at the ends of the stimulus series.

The retraining and retesting gradients represent combined data for the two days on which each of these procedures was carried out. The response rate recovered quickly upon retraining, and these gradients lie above the VI training baselines. There is a small remaining depression at or near 570 Mμ for both groups. The re-extinction gradients are, of course, at a much lower level and the residual depression found during retraining in the area of 570 Mμ is still evident.
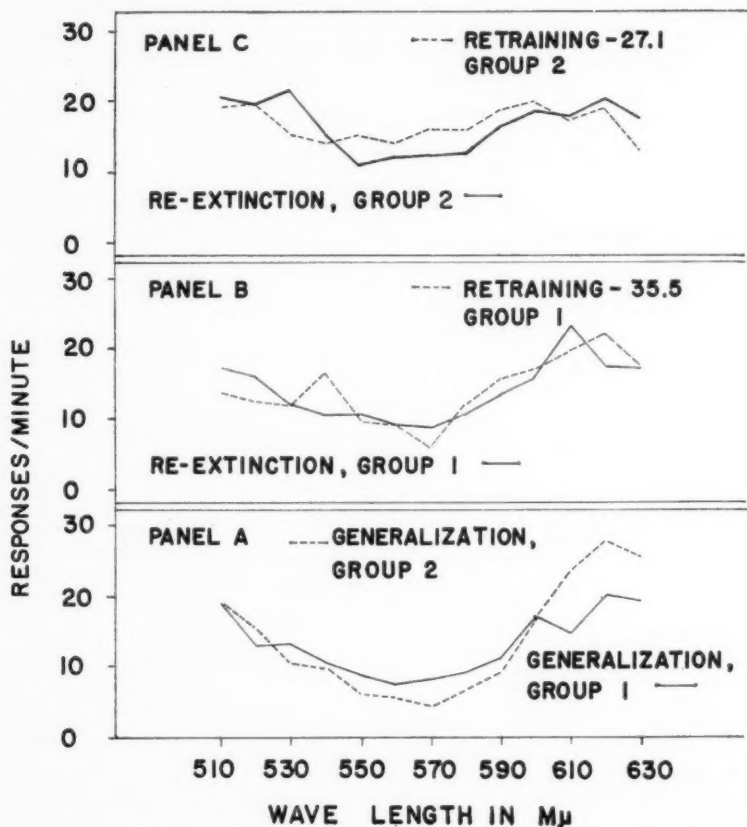
Fig. 4. Some selected comparisons. In Panel A the combined gradients for both days of testing can be compared for Groups 1 and 2. In Panels B and C retraining and retesting (re-extinction) gradients can be compared for both groups after the retraining gradients have been brought to the same mean level as the retesting gradients by subtracting an appropriate constant from each value on the former.

As indicated in Fig. 3, the gradients obtained without specific extinction at 570 Mμ from the Control Group are flat, reflecting equal rates of extinction at all stimulus values during the generalization test. This eliminates the possibility that the depression around S^e for the gradients from the experimental groups is due entirely or in part to the testing procedure rather than specific extinction at 570 Mu. It could be argued that in the course of the test under extinction, a gradient of extinction would be generated around each stimulus value, and that these individual gradients could summate, leading to greater decrement in the middle of the stimulus series. It is clear that such a summation does not occur, lending support to recent studies that have investigated summation of generalization gradients of acquisition (Kalish and Guttman, 1957, 1958).

It is of interest to make a few selected comparisons between gradients; these are presented in Fig. 4. In panel A the combined generalization gradients for both days of testing are compared for Groups 1 and 2. The gradient for Group 2 looks steeper than that for Group 1, which would reflect the greater length of the period of extinction at 570 Mμ for Group 2. But neither the slope nor the overall response level is shown to be significantly different for the two groups by a mixed analysis of variance where Ss and stimuli are within-groups variables and amount of specific extinction is the between-groups variable (Lindquist Type 1, 1953).

Another comparison of interest is between the retraining and the retesting gradients. This is facilitated in panels B and C of Fig. 4, where a constant value has been subtracted from each point on the retraining gradients, so that they are brought to the same level as the retesting gradients. While the gradients flatten out rapidly during retraining, it is quite conceivable that the depression associated around S^e would reappear during retesting, since this condition is similar to the first generalization test. This does not seem to be the case, however, as the retesting gradients differ from the retraining gradients only in overall response level but not in slope. In this respect the retesting procedure had just the same effect on the retraining gradient as the general extinction for the Control Group had on the VI baseline.

DISCUSSION

There is no doubt that reliable gradients of extinction were obtained with the present method. They are not artifacts of averaging, since clear gradients can be seen in the individual data for 8 out of 12 birds. It is evident also that these gradients are not of the same form as (i.e. inversions of) the Guttman and Kalish gradients of acquisition, which are steep near the CS and flatten out as the response level approaches the operant rate. In the present study the greatest decrement (i.e. the steepest rise in response level) occurred not near the S^e but at more remote values.

It could be concluded that extinction (or inhibition) generalizes more broadly than acquisition (or excitation), but some words of caution are necessary. During the initial stages of generalization testing, the response rates at the ends of the gradient were equal to the rates during acquisition, thus reaching the pre-extinction level within 50 or 60 Mμ of the S^e value. In studies of the gradient of acquisition, a similar span is found between the CS and the point at which the response level reaches the operant rate. Thus the generalization of extinction, though flatter near the S^e, does not extend further in terms of total range.

A second point regards the pre-test experience of the Ss. Previous to a test for the gradient of acquisition, only the training value is normally presented; S is not presented with the range which is to be used on the test. The procedure for obtaining a gradient of extinction involves the presentation of a large range of equally reinforced stimuli during the acquisition period. If this reduces any tendency to respond differentially among stimuli, the shallower gradient of extinction is not surprising. This difficulty is inherent in any study like the present one using the traditional design for obtaining gradients of extinction.

To obtain truly comparable gradients of acquisition and extinction, it is necessary to make the pre-extinction experience for the Ss comparable. Kling approached this by training Ss on no more than two stimuli, and then extinguishing on one. But, as noted above, he used independent groups to estimate the gradient of extinction, and he did not cite any comparable studies of the gradient of acquisition. An even closer approach to the design of the typical acquisition study is that of Honig, Thomas, and Guttman (1959), who trained pigeons at one stimulus only (550 Mμ), gave extinction at a different value (570 Mμ) and followed this with a generalization test very similar to the one used in the present study. The post-extinction gradient was compared to that of a control group which had received acquisition training at 550 Mμ but no extinction at 570 Mμ.

The post-extinction gradient was found to be similar to the control gradient except that it was reduced by a constant proportion; the reduction was neither specific to nor centered around the S^e. It appears, then, that making the pre-test stimulus exposure similar to the gradient of acquisition studies fails to produce any gradient of extinction in the usual sense. In view of this, it becomes inappropriate to ask why the gradient of extinction is flatter than that of acquisition; the pertinent question is why the design used in the present study should yield a gradient at all while that of Honig, Thomas, and Guttman fails to do so. While this is not the place for a detailed theoretical analysis, the basis for an explanation may be outlined informally.

The first assumption is that there is *no* primary gradient of extinction comparable to the gradient of acquisition. Instead, repetition of a

response without reinforcement reduces the learned habit as a whole, irrespective of the particular stimulus present during extinction.[2] The second assumption is that the generalized response strength comprising the gradient of acquisition at different points is simply part of the acquired habit. These assumptions would permit the explanation of the Honig, Thomas, and Guttman results, where extinction at 570 Mμ reduced the gradient around 550 Mμ simply as if the extinction had occurred at the training value.

For the explanation of the present results, it is necessary to assume further that the response strength at a given stimulus (e.g. 570 Mμ) results not only from the training given at that value, but also from the generalization of training received at other values, e.g. 560 Mμ, 580 Mμ, and so forth.[3] It would seem reasonable that each of these separate habits contributes to response strength at a given training value to a degree inversely related to the difference between the two training values. Finally we may assume that the degree of the reduction of each contributing habit due to extinction at a given value (570 Mμ) is directly related to the strength of its contribution at that value. In such a case, the empirical gradients of extinction obtained here are to be expected. But it should be kept in mind that according to the present analysis, such gradients are not "primary" generalization gradients at all, but result from complex interactions among differentially reduced gradients of acquisition.

To give direct support to such a theory, there should be a specific mathematical formulation of assumptions dealing with the forms of the gradients, the interaction of response strengths, and so forth. At this time, such a formulation has not been worked out, nor are there data which would provide it independent of the results of the present study. The "explanation" of a given set of results by the combination of hypothetical parameters is usually possible, but it does not give much independent or empirical support to the theoretical formulation. The rough qualitative analysis presented above must suffice here, and the author can do little more in its support than to point out that all previous studies of the gradient of extinction, even that of Kling, have involved acquisition with more than one stimulus previous to extinction and testing. The derivation of their empirically obtained gradients from differentially reduced gradients of acquisition is therefore within the scope of the proposed analysis.

[2] This assumption is applicable only to a situation involving continuous extinction; the alternation of positive and negative conditions in discrimination learning produces a clear response decrement around Se (or S-).

[3] This assumption implies a summation of response strengths that has been questioned by Kalish and Guttman (1957, 1958). It should be noted that they studied the summation of generalized response strengths at values other than the training stimuli, while our assumption refers to summation among training values. Furthermore, they gave much more training at each of their two or three training values than we did at each of 13 values. Summation may occur primarily with limited training at a large number of values. It is interesting that Blough (1959) has suggested a similar summation to explain brightness preferences of pigeons given equal reinforcement at a number of brightness values.

SUMMARY

Generalization gradients of extinction on the spectral continuum
were determined for pigeons by training them to peck at 13 different
spectral values ranging from 510 to 630 Mμ, extinguishing at one value
(570 Mμ), and then testing the response rate under extinction at all
the original training values. Clear decremental gradients were obtained
around 570 Mμ. The decrement was largely abolished by retraining at
all stimulus values, and did not reappear when retesting was carried
out without further extinction at 570 Mμ. In the discussion, this study
is compared to a previous one where training at a single stimulus value
followed by extinction at a different value failed to produce a decre-
mental gradient of extinction.

An informal argument is presented in order to explain the differ-
ence in results. This assumes that the present gradients are not pri-
mary gradients of extinction at all, but shows that they could result from
a combination of gradients of acquisition around the different training
values which we differentially reduced by the extinction procedure.

REFERENCES

BASS, M. J., and HULL, C. L. Irradiation of a tactile conditional reflex in man.
    J. comp. Psychol., 1934, 17, 47-65.

BLOUGH, D. S. Generalization and preference on a stimulus-intensity continuum.
    J. exp. Anal. Beh., 1959, 2, 307-317.

GUTTMAN, N., and KALISH, H. I. Discriminability and stimulus generalization.
    J. exp. Psychol., 1956, 51, 79-88.

HANSON, H. M. Effects of discrimination training on stimulus generalization.
    J. exp. Psychol., 1959, 58, 321-334.

HONIG, W. K., THOMAS, D. R., and GUTTMAN, N. Differential effects of con-
    tinuous extinction and discrimination training on the generalization gradi-
    ent. *J. exp. Psychol.*, 1959, 58, 145-152.

HOVLAND, C. I. The generalization of conditioned responses: I. The sensory
    generalization of conditioned responses with varying frequencies of tone.
    J. gen. Psychol., 1937, 17, 125-148.

HULL, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1953.

KALISH, H. I., and GUTTMAN, N. Stimulus generalization after equal training on
    two stimuli. *J. exp. Psychol.*, 1957, 53, 139-144.

KALISH, H. I., and GUTTMAN, N. Stimulus generalization after training on three
    stimuli: A test of the summation hypothesis. *J. exp. Psychol.*, 1959,
    57, 268-272.

KLING, J. W. Generalization of extinction of an instrumental response to stimuli
    varying in the size dimension. *J. exp. Psychol.*, 1952, 44, 339-346.

PAVLOV, I. P. *Conditional reflexes.* Trans. G. V. Anrep. New York: Oxford
    Univer. Press, 1927.

SPENCE, K. W. The differential response in animals to stimuli varying within a
    single dimension. *Psychol. Rev.*, 1937, 44, 430-444.

A REEXAMINATION OF THE TWO KINDS OF SCIENTIFIC CONJECTURE[1]

WALTER S. TURNER

University of California at Berkeley

Flee the precepts of those speculators whose reasonings are not confirmed by experience.—Leonardo da Vinci (1452-1519)

Most statements about the role of theory in science apply to only one of the two kinds of scientific conjecture: i.e., either to conjectures based on abstraction from sensory observations, or to conjectures that postulate imagined properties. Because such statements usually fail to specify which kind the author means, they are usually ambiguous and confusing. This paper urges more general recognition of the distinction between the two types, and specification of the type referred to whenever a statement does not apply to both. It also asserts that conjectures about imagined properties can never be empirically supported, and, on the whole, do more harm than good. It particularly deplores the common practices of debating the probability of such conjectures and making them the foci of experiments.

THE NEED FOR THE DISTINCTION

Petrus Ramus, after disparaging the work of Copernicus, once offered his professorship as a prize for an astronomy constructed without hypotheses. When Kepler replied that Copernicus had already produced such an astronomy, it was clear that Kepler and Ramus meant different things by "hypotheses." When Bacon, who advocated the hypothetical method (Blake, *et al.*, 1960, pp. 54, 71-74), spoke out repeatedly against the use of "theories," it is plain that he did not condemn all types of theory. When Newton advocated drawing tentative conclusions by induction from experiments and observations, and in the same paragraph said, "hypotheses are not to be regarded in experimental philosophy" (*Opticks*, p. 380) he too must have meant to condemn only a particular kind of conjecture. What kind was it?

Whenever philosophers and historians of science have dichotomized conjectures into good and bad, the line between them has been drawn in essentially the same way. The approved type has been based on abstraction: i.e., a class of events is endowed with properties abstracted from observations of a sample of that class. The disapproved type has been based on imagination: i.e., both the sample and the class

are endowed with imagined properties, because their presence would produce the observed phenomena.

Bacon condemned the second type for its fantastic suppositions to which nothing in reality corresponds (1620b, p. 519); "God forbid that we should give out a dream of our own imagination for a pattern of the world" (1620a, p. 451). Newton—despite his occasional use of them (usually apologetically, in personal letters, or disguised as "queries")—thought conjectures of this type had no place in experimental science (Cajori, 1934, p. 676). Jevons credited all the major successes of science to the first (abstractive) type (1874, pp. 137, 431), and feared that Type 2 (imaginative) conjectures might divert science from the search for empirical laws (p. 155). Duhem (1914, p. 43) found all that is good in a theory in the abstractive ("descriptive") part, to which the imaginative ("explanatory") part attaches itself like a parasite. Skinner (1950) approves Type 1 theories (p. 215f) but finds Type 2 theories unnecessary, of doubtful utility, and often harmful (p. 193ff).

Which kind of conjecture is referred to in Russell's "argument in favor of a theory is always . . . invalid" (1927, p. 194)? Or Dallenbach's "Doing something, without a theory, is . . . mere busy work" (1953, p. 38f)? Or Lewin's "Theories are unavoidable" (1954, p. 920)? Or Koch's "Psychology is not ready for high-order theory" (1956, p. 43)? Or Beveridge's "hypothesis is the principal intellectual instrument in research" (1957, p. 52)? Or Sidman's contention that "theory" is not needed to organize data (1960, p. 13ff)? Are these men all referring to the same type of conjecture? Can we understand them without knowing which type they mean?

THE TWO TYPES DISTINGUISHED

"Conjecture" in this paper means "opinion held, with much or little confidence, but without enough evidence for certainty." It includes both theory and hypothesis—here used interchangeably to denote more or less formal and considered conjectures. It also includes hunch, guess, surmise, and (at the other end of the strength-of-evidence continuum) scientific law—since scientific laws are never absolutely certain.

The historical abstractive-imaginative dichotomy of conjectures is delineated and illustrated by Dingle (1931, p. 22f) and Woodger (1956, p. 17ff). It has been applied by psychologists both to conjectures (e.g., Pratt, 1945; O'Neil, 1953; Farrell, 1955) and to concepts (MacCorquodale & Meehl, 1948; Maze, 1954; Rozeboom, 1956); but it is not sufficiently understood.

Abstraction is the process whereby a property is seen to be common to two or more observations. When possession of an abstracted property is stated as the criterion of membership in a class, the result is a definition or classification. But if an abstracted property is conjectured to characterize a class *not defined in terms of that property*, the result is an abstractive (Type 1) conjecture.

To illustrate: if we observe A and B with IQ's of 160, and set up a class consisting of all persons with IQ's of 160 or more, we have simply defined a class—to which we can give any name we choose: e.g., "genius." But if, from the same observations, we abstract the additional common property of leanness, and then—without adding this property to our definition—we conjecture that all or most geniuses are lean, we have a Type 1 conjecture.

Examples of properties abstractable from sensory observations are: size, shape, motion, speed, distance, density, transparency, ductility, solubility, conductivity, chemical valence, mathematical relations, overt behavior probabilities, and latency, frequency, and amplitude of overt responses.

In a Type 1 conjecture the property attributed to a class must be abstractable from *past sensory observations of members of that class.*[2] If the attributed property is abstractable only from observations of members of a class other than the one named in the conjecture, the conjecture is Type 2. Such conjectures, however, are rarely published; for if the attributed property has been observed elsewhere, the theorist will almost certainly try to observe it in a sample of the named class. If he succeeds, the conjecture (by definition) becomes Type 1. If he fails, he will probably not publish his conjecture.

Most published Type 2 conjectures, therefore, attribute properties (hypothetical constructs) which are not abstractable from *any* past sensory observations—such as God, id, ego, self, will, desire, purpose, identification, repression, Kohler's cortical field, Hebb's cell assembly, Hull's constructs in general, and traits or feelings such as aggressiveness (as distinct from aggressive *behavior*).

Three points in the Type 1–Type 2 distinction need emphasis. First, a conjecture is Type 1 if the properties it attributes *can be* abstracted (regardless of how they were originally derived) from past sensory observations of members of the named class. Thus a Type 2 conjecture can become Type 1—which transformation, in fact, automatically occurs the moment the properties attributed by a Type 2 conjecture to a class are observed in a sample of that class. Type 1 conjectures can never become Type 2; but many conjectures have changed from Type 2 to Type 1.

Second, "abstractable" does not mean "inferable." By observing moving bodies we can abstract the property of motion. By observing people with loud voices and red faces we can abstract the properties of voice-loudness and face-redness—but *not* the property of anger. If we read anger into our Ss we are just inferring it as a likely cause of what we observed. Abstraction, however, is not necessarily a mechanical act. It may involve ingenious discovery of properties far from apparent on the surface—as in the case of Kepler's much-discussed discovery that planets have elliptical orbits.

---

2 Examples of Type 1 conjectures are: Jost's laws, Marbe's law, Ebbinghaus's forgetting curve, Osgood's laws of transfer and retroaction, Berg's deviation hypothesis, conjectures as to correlations between test scores and overt criterion variables, and most of the major theories in physical science.

Kepler's conjecture was clearly abstractive and Type 1. Previous astronomical observations had given numerous loci of the several planets at specified times. For each planet these loci lay in a roughly elliptical path in three-dimensional space. Their collective ellipticality was not easy to see; but it was in the loci themselves—Kepler did not put it there. By trial and error he discovered the ellipticality, and then conjectured that this property of the observed points in the orbits was also present in the unobserved points.

Third, "sensory observations" does not imply "unaided senses." Observation via telescope or microscope is still sensory observation. To be sure, we do not actually observe, for example, blood pressure or GSR by watching movements of a needle on a dial. But these properties are abstractable from a variety of measurement operations which have been performed in the past, and not just from the particular type of operation used in the given instance.

It may be argued that if the property of blood pressure is abstractable from a variety of measuring operations, then the property of anxiety should be abstractable from a different set of operations. But the correlation between measurements of blood pressure obtained by different operations is very high—as high as could be expected from imperfect measuring devices; while the correlation between different measures of anxiety is so low as to suggest that different variables are being measured.[3]

It is theoretically possible, of course, to devise quite different instruments for measuring "anxiety" which will agree with each other very reliably (cf. Campbell & Fiske, 1959, p. 101). Conjectures about a property abstracted from observation with such instruments (or about *scores* on any measuring instrument) would be technically Type 1. But the moment the theorist uses any ill-defined, vernacular term like "anxiety" in his conjecture, he is conveying *to his public* the idea of a property which is *not* abstractable from sensory observations.

TYPE 2 CONJECTURES AND VERIFIED PREDICTIONS

About 1300 A.D., Bernard of Verdun anticipated the argument of the modern believers in the empirical confirmability of Type 2 conjectures. After noting that all the astronomical predictions deduced from Ptolemy's (later-discredited) Type 2 conjectures had been verified, he argued that "this would certainly not have happened if the point of departure of these deductions had been false" (Blake, *et al.*, 1960, p. 24).

In 1495 Capuano stated the same argument even more incautiously. He argued that an exact accord between empirical results and deductions from Ptolemy's hypotheses "demonstrates with certainty" the truth of the hypotheses (Ibid.).

Half a century later Copernicus claimed that all known astronom-

[3] If the question then be asked, "How high a correlation is high?" there can be no satisfactory theoretical answer. But we are not kept from easily distinguishing bearded from beardless men by our inability to state how many hairs make a beard.

ical observations could be predicted by his theoretical system, and that any change in it would destroy its predictive power. From this he seems to have concluded that his system was the only possible one, and therefore demonstrably true. Kepler also viewed the Copernican theory as demonstrated with certainty by its "consistent explanation of so many phenomena" (Blake, *et al.*, 1960, pp. 27, 38).

In 1581 Clavius used Corpernicus's argument to support Ptolemy's Type 2 theory of solid eccentric and epicyclic spheres. He said no known theory (including Copernicus's) can predict the observed phenomena without positing such spheres; therefore their existence is highly probable (Blake, *et al.*, 1960, p. 32).

In 1644 Descartes defended his (since-discredited) Type 2 conjectures (as to the cause of certain phenomena) by the same argument of confirmation by consequences. He wrote: "when it seemed to me impossible to find in the whole extent of nature any other cause capable of producing them, I believed that they had infallibly been so produced" (Blake, *et al.*, 1960, p. 94.).

The fallacy in all these claims was nailed by Newton. Agreement of observed facts with deductions from a Type 2 conjecture does *not* support the conjecture because the same facts might also be deduced from numerous other conjectures. (Brown, 1950, p. 94f; Blake, *et al.*, 1960, pp. 126-135). Such empirical verification merely lets us classify the conjecture as one of the many which are compatible with the facts.

Philosophers and historians of science since Newton have generally agreed that Type 2 conjectures can *not* be confirmed or supported by verifying their empirical predictions. The scientific (or Newtonian or hypothetico-deductive) method applies *only* to Type 1 conjectures. Indeed this follows automatically from the classical position that empirical validation can come only through induction; for the basis of induction—observation of the attributed property in a sample of the specified class (Mill, 1874, pp. 210, 221, 223; Peirce, 1878, pp. 272, 277, 280)—is by definition absent in Type 2 conjectures.

Whewell (1858) insisted that in every induction some unifying conception is invented or superinduced upon the facts (p. 74) to describe their "mode of connexion" (p. 202). His "invented conceptions," however, are *abstracted relations;* once discovered, they are seen to have been present in the facts from the start (Blake, *et al.*, 1960, p. 202). Hence his rule for testing hypotheses is to see whether the facts have the relation supposed by the hypothesis (p. 67). Thus for Whewell, as for Mill and Herschel (1842), the only conjectures empirically supportable are those stating abstractable empirical properties. He called Type 2 conjectures "superfluous and irrelevant" (p. 66f).

Despite the three-century consensus that Type 2 conjectures cannot be empirically supported, scientists have never stopped making them. Recently, moreover, many psychologists and a few logicians have raised their voices against the classical position and revived the me-

dieval fallacy that Type 2 conjectures can be confirmed by verifying their predictions. Hull stated in 1935 that indirect verification occurs when a deduction from a combination of postulates is observationally confirmed (p. 510), and in 1937, that verification of its theorems makes the system (of Type 2 conjectures) "probably true" (p. 5). The falsity of these assertions was quickly pointed out (Adams, 1937; Miller, 1939), but without visible effect on Hull or his followers.

In 1949 Kneale (p. 362) declared that evidence for the consequences of a Type 2 conjecture ("transcendent hypothesis") is evidence for the conjecture. He admitted that such verification cannot be considered ordinary induction (since there are no observed instances from which to infer); but he proposed to legitimize the procedure by calling it "secondary induction." The same year, however, Reichenbach (1949, p. 432) noted that any procedure relying on indirect confirmation contains the elementary logical fallacy of supposing that if A implies B, then A is probable if B is given.

In 1952 (p. 35f) Hempel argued that deducing enough empirical predictions from a network of Type 2 conjectures makes the whole system testable, and that confirmation of the predictions strengthens the parent conjectures. Similarly Feigl (1959, p. 127) states that merely embedding theories in a deductive network "anchored only in a few places in the facts" makes possible their empirical confirmation.

Indirect confirmation of a network of Type 2 conjectures, however, is still indirect confirmation—and still fallacious. We may make 20 conjectures as to how X would feel if he were possessed by demons, and deduce 100 predictions as to how he would act; but verifying all 100 predictions would be *no evidence whatever* that X was possessed by demons—since the same behavior could be deduced from an infinity of other hypotheses. To argue "the truth of an hypothesis by affirming the consequents is always a fallacy no matter how many consequents are affirmed" (Johnson, 1954, p. 727). When a consequent is affirmed, "it does not follow that the probability of the theory is increased" (McGuigan, 1956, p. 98).[4]

Feigl (Ibid.) also says, without adducing evidence, that indirect confirmation of theories is "the essential feature of all sciences which advance beyond the observational, fact-collecting ('botanizing') stage." It is true that many sciences fabricate Type 2 conjectures and test their empirical deductions; but no Type 2 conjecture was ever supported by empirical evidence in *any* science. Such conjectures, says Russell (1940,

---

[4] Some writers (e.g., McGuigan, 1956) have tried to justify indirect confirmation on the basis of Bayes's rule. But Reichenbach (1949) points out that Bayes's rule is applicable only when much more is known than the occurrence of the consequences of a conjecture (p. 432); we must know its "antecedent probability" and the probability that other conjectures might entail the same consequences (p. 95). Both probabilities are unknown in the case of scientific conjectures. Polya (1954, v. 1, p. 22, v. 2, p. 3f) speaks of inductively supporting a mathematical conjecture by verifying its "consequences," but his examples show that by "consequences" he means "observed instances." Thus to support Euler's conjecture that $8n+3=x^2+2p$ he simply substitutes for n the integers from 1 to 10, one at a time, and shows that in all ten *instances* the equation holds.

p. 381ff), "cannot be rendered probable or improbable by any empirical evidence . . . there can never be any evidence in their favor."

Indirect confirmation underlies the concept of "construct validity." This concept, introduced by an APA Committee (1954, p. 14) and much praised since (Cronbach & Meehl, 1955; Jenkins & Lykken, 1957, p. 81; Loevinger, 1957, p. 636), has unfortunately become part of the vocabulary of psychology (cf. Bechtoldt, 1960). Construct validation starts with a hypothetical construct: for example, an invisible little man as the soundmaker in our radio. We might then predict that since the man would periodically sleep, the radio should be periodically silent—probably late at night. If we then find that our radio is in fact silent at night from 2 until 5, the construct validationists call this confirmed prediction "indirect evidence" that the hypothetical little man does exist. The confirmed prediction does not, however, really support the little man construct, since the same results are predictable from an infinity[5] of other constructs.

TYPE 2 CONJECTURES AND SCIENTIFIC RESEARCH

A great many writers who do *not* claim that Type 2 conjectures are confirmable, still argue that they generate Type 1 hypotheses, lead to experimentation, and so aid the discovery of empirical laws. This position, first taken by Whewell, has since been uncritically echoed by hundreds of writers—e.g., Russell (1918, p. 158), Brown (1950, pp. 94, 112f), Conant (1951), Marx (1951, p. 239f), Kendler (1952), Zubin (1952), Kessen & Kimble (1952), Braithwaite (1953, pp. 267, 272), Sarbin (1954, p. 245), Woodger (1956, pp. 19f, 35f), and Beveridge (1957, p. 46).

Conant (1951, p. 25f), for example, says the usefulness of all scientific hypotheses (Type 1 or 2) lies in the research they generate. Woodger (1956, p. 78) says the chief role of Type 2 hypotheses is to "keep researchers busy." While much, if not most, psychological experimentation is oriented toward Type 2 theories, no one has ever shown that these experimenters would have run out of experiments, or been any less "busy" if they did not have these theories. Much useful psychological research goes on without reference to Type 2 conjectures (in fact, an entire journal is devoted to such experiments), and the amount that could be profitably conducted is limited only by the number of qualified researchers available. Clearly, Type 2 conjectures are *not* necessary to keep researchers busy.

Sidman (1960, pp. 4-40) notes the many psychological experiments made for other reasons than to test Type 2 hypotheses. He lists experiments to establish the existence of a behavioral phenomenon, to

[5] If we grant the possibility that the true cause of any phenomenon may be some unthought-of alternative to the cause we have conjectured, we thereby confess that we don't know how many conceivable alternatives we may have overlooked. If we don't know the number of possibly overlooked alternatives, and are unable to specify a finite number which it can't possibly exceed, we are forced to regard the total number of possible alternatives as infinite.

explore the conditions under which a phenomenon occurs, to try out a new technique or method, and—often the most productive—to answer the question, "I wonder what will happen if . . ." He finds the importance of data unrelated to the sophistication of the hypothesis that may have generated the experiment, and adds that few psychologists "would deny that the most interesting behavioral phenomena have not even been touched by the most rigorous present-day theories."

Type 2 conjectures clearly impede progress when they cause substitution of less productive for more productive research. Underwood (1957, p. 185) thinks much effort has been wasted in "bitsy-type research," designed to test theoretical disagreements, which has failed to establish reliable relations between observable variables. Skinner (1950, p. 194) likewise thinks much useless experimentation comes from Type 2 theories, and much energy and skill are absorbed by them. Sidman (1960, p. 4) notes that research by men whose curiosity about nature is less than their curiosity about the accuracy of their conjectures "can result in the piling up of trivia upon trivia." Davis (1953, p. 9) notes that when a Type 2 theorist dreams up a physiological referent for his construct, as Krech (1950) would have him do, he may send physiologists on a fool's errand in quest of his possibly non-existent referent.

Since Type 2 conjectures have not demonstrably increased the total amount of research, the research they generate must be considered a diversion, not an addition. If psychologists stopped doing (and discussing!) experiments involving Type 2 conjectures, they would have much more time for purely empirical research. Furthermore, it seems likely that scientists obtain knowledge of empirical relations faster by looking for it directly, than by looking primarily for something else (i.e., confirmation of Type 2 conjectures). Therefore, by diverting us from more productive lines of inquiry, Type 2 conjectures may have seriously reduced the total harvest from our research.

Beveridge (1957, p. 51) notes that Type 2 conjectures may, for long periods, slow the advance of science: witness the phlogiston doctrine and the theory that every metal contains mercury. Harlow (1953, pp. 24, 27) once felt that progress in his field was slowed by misdirected and unfortunate channelization of effort due to psychologists' preoccupation with Hull's "drive" theory. Dallenbach says, in a letter to the author:

> I think that Type 2 conjectures are useless and, moreover, downright harmful. Particularly is this true in psychology in which we have a plethora of theories of this type which distract the attention and engage the efforts of so many of the real experimenters we have. Hull's theoretical constructs are cases in point . . . We would be far ahead in our knowledge about learning if Hull's constructs had been ignored.

Type 2 theories are often said to "lead to" the discovery of empirical laws. Whewell (1858, p. 83), for example, said that even false hypotheses, if they bind together facts which without them are loose and detached, may "lead the way" to the true empirical relationship between them. He cited several Type 2 theories which were later *followed* by Type 1 laws, but he failed to show that the former *led the way* to the latter; they may actually have blocked the way.

Brown (1950, p. 112f) cites the electromagnetic theory of light as being originally based on the now-discarded ether phyothesis. But "based on" does not equal "suggested by"; and since the ether hypothesis antedated by two centuries the electromagnetic theory, their timing does not suggest that one generated the other. On the contrary, the evidence suggests that the latter theory was conceived because the accumulation of knowledge (Type 1 data) about electrical phenomena paved the way for it. Beveridge (1957, pp. 41-46) cites several scientific hypotheses which led to important discoveries, but none were Type 2.

Woodger (1956, pp. 20, 36) says Type 2 conjectures often lead us to manipulate things in ways we would probably not have thought of without them and suggest hitherto unsuspected directions of investigation. When the manipulations and investigations turn up an empirical law, it is easy to say, "There now! You see how useful the Type 2 theory turned out to be!" But how many blind alleys did the theory previously lead scientists to explore? What explorations and discoveries would have been made without the theory? How many other Type 2 theories engaged scientists' attention during the same period without producing worthwhile results?

Braithwaite (1953, p. 272) says Type 2 hypotheses have enabled scientists to make many correct predictions; but so do Type 1 hypotheses. Duhem (1914, p. 43) says it is precisely the Type 2 component in physical theories that causes most of their predictive failures. The evidence that Type 2 conjectures lead to discovery of empirical laws usually reduces (as in the examples cited by Whewell and Brown) to the simple fact that such discoveries have sometimes *followed* Type 2 conjectures—possibly after a lag of centuries. Lacking any evidence of casual connection, we have no more reason to think that the conjectures speeded up the eventual discoveries than that they retarded them.

Finally, commitment to Type 2 theories prejudices scientists and tends to blind them to research results incompatible with their theories. Discrepancies between facts and Type 1 theories are less likely to exist and persist, since Type 1 theories derive from facts and depend solely on facts for support. Long ago Rankine (1855, p. 212) observed that received Type 2 hypothesis tend to lead scientists:

. . . to explain away, or set aside, facts inconsistent with these

hypotheses, which facts, rightly appreciated, would have formed the basis of true theories. Thus, the fact of the production of heat by friction . . . was long neglected because inconsistent with the hypothesis of calorie; and the fact of the production of cold by electric currents . . . was, from inconsistency with prevalent assumptions respecting the so-called 'electric fluid', by some regarded as a thing to be explained away, and by others as a delusion.

TYPE 2 CONJECTURES AND UNDERSTANDING

If Type 2 conjectures can never receive empirical support, and more often hurt than help research, are they of any use whatever? Many scientists (and more non-scientists) would reply that only Type 2 conjectures provide an explanation of nature, i.e., enable us to understand nature. For others, however, such conjectures are, by definition, unable to explain anything, or to contribute directly to our understanding. This conflict of opinion reflects different conceptions of the meanings of "explanation" and "understanding."

To explain means "to make understandable." For some, psychological phenomena are understandable only when we know the observable conditions of which they are a function; hence only Type 1 generalizations can explain. For others, psychological phenomena are understandable only when an internal, unobservable cause has been assigned. For them, only Type 2 conjectures can explain.

For the latter, Type 1 generalizations describe rather than explain; even the highest-level empirical laws, for these people, contribute no insight or understanding of natural phenomena, because they do not reveal inner causes or forces behind the phenomena. For them, only the unseen can explain the seen (Peirce, 1878, p. 477; Duhem, 1914, p. 3f); explanation must reveal another world behind the scenes of the observable (Woodger, 1957, p. 20). For example, the physicist Gamow (1958, p. 6) says:

> The role of theoretical science is to find the hidden interrelations between the empirical laws and to interpret them in the light of certain hypothetical assumptions concerning the internal structure of matter and various material objects which are not subject to direct observation.

Type 2 explanations, as has been shown, can never have any measurable or estimable probability of being true. If they do not contradict known facts, there is merely a *possibility* of their being true. The *probability* of one Type 2 explanation can never be said to be greater than that of another. Since they are essentially fictional models, a decision to accept one of them has "no truth character." Our choice "cannot be evaluated as being right or wrong. It is purely a matter of personal taste" (Kendler, 1952, p. 276).

Clearly, then, those who seek revelations of a world behind the phenomena are no better off with Type 2 than with Type 1 explanations; for Type 2 explanations do not *reveal* such a world—they merely *fabricate* one. The popularity of such explanations may well be largely due to the common failure to recognize their essentially literary or fictional character.

Natural science is essentially the systematic seeking of empirical truth. Type 2 explanations, therefore, have no place in it—no matter how thoroughly they now infest it, no matter how useful they may be in philosophy, and no matter how subjectively satisfying they may be. Their own probability of being true can be neither measured nor estimated, and, as has been shown, they more often slow than speed the discovery of truth. Skinner (1953, p. 24) calls such explanations "prescientific":

> The study of any subject begins in the realm of superstition. The fanciful explanation precedes the valid. Astronomy began as astrology; chemistry as alchemy. The field of behavior had, and still has, its astrologers and alchemists. A long history of prescientific explanation furnishes us with a fantastic array of causes which have no function other than to supply spurious answers to questions which must otherwise go unanswered in the early stages of a science.

Rankine (1855, p. 213) considered Type 2 conjectures necessary in molecular physics (which was not his field) to reduce the data "to simplicity and order" as a preliminary to the framing of Type 1 conjectures. Marx (1951, p. 245) calls hypothetical constructs a "temporary expedient in the development of sound psychological theory." But the Rankine-Marx position is just the opposite of that of the modern English physicist, Whittaker (1952, p. 48). He believes that development of Type 1 conjectures must come first: "The formulae of reflection and refraction must be known before Huygens can devise his Principle to explain them."

Woodger (1956, p. 20) suggests that Type 2 conjectures increase our understanding by enabling us to classify data which would otherwise be unrelated and unmanageable. Type 2 conjectures, however, are not needed for this purpose, since Type 1 generalizations are preeminently fitted for the job. As Sidman (1960, p. 15) says, methodical research in any area by an alert observer will inevitably bring out interrelations among the phenomena in that area, in the form of similarities among the relevant variables. He considers systematization of the data in this way, rather than by Type 2 theories, "a vital prerequisite" for the development of our understanding of behavior.

For a great many, if not most, modern writers, a scientific explanation is *by definition* a Type 1 generalization. Thus, for Bridgman (1927, p. 47) and Pratt (1945, p. 266) a scientific explanation is sim-

ply a statement of a correlation between observed events. For Berg-
man & Spence (1941, p. 4) it is inductive generalization of empirical
laws—i.e., of functional relations between variables. For Skinner it
states the variables of which behavior is a function (1957, p. 10); ref-
erence to unobservables is in no sense an explanation, and adds nothing
to a functional account (1953, p. 144). For Guthrie (1933, pp. 125,
130) it is showing an event to be an instance of an empirical law, with-
out referring to unobservables.

Bergmann (1957, pp. 75-84) says that description states particular
facts, while explanation states laws. But in a more general sense of
description, explanation is a species of it. Schlick (1949, p. 20) calls
explanation of nature "description of nature by means of laws." Geldard
(1939, p. 414) finds that explanations consist simply of further descrip-
tions.

It is thus clear that Type 1 generalizations are the scientific way to
explain nature. Type 2 conjectures can provide only literary, fictional,
nonscientific explanations which add nothing to our knowledge.

HYPOTHETICAL CONSTRUCTS

In psychology a hypothetical construct is a process, property, or
entity that is held to exist in an organism, although it is not abstract-
able from past sensory observations of the class of organism involved.
Hypothetical constructs always entail Type 2 conjectures; but the at-
tention they have received in the literature entitles them to separate
consideration. Boring (1953, p. 172) calls them the stuff of which
science is made.

Three features of hypothetical constructs need emphasis:

1. Certain overt behavior may be commonly assumed to accom-
pany or reflect the construct; but since the relation is only assumed,
the construct is not abstractable from the behavior. Hunger is not ab-
stractable from salivation, nor fear from avoidance.

2. The construct may or may not be "operationally defined," *as
this term is used in psychology.* Constructs are operationally defined
in psychology by stating the operations used to distinguish them (Tol-
man, 1932; Stevens, 1939; Ritchie, 1954; Torgerson, 1958; English &
English, 1958). Thus, while the operations distinguish the constructs,
they are *not synonymous* with the constructs. The construct is still
conceived as an unobservable process or entity, located inside the or-
ganism. For example: in an experiment using GSR as the sole measure
of changes in "drive state," the experimenter would never say that
drive *is* GSR and nothing more. The operational definition merely
states the conditions under which the presence of the unobserved event
is to be inferred; the inferred event remains a hypothetical construct.

3. The construct may or may not be abstractable *in principle,* as
MacCorquodale & Meehl (1948, p. 104) point out. It may be physical,
physiological, or psychic, so long as it is not abstractable from past

sensory observations. To require a hypothetical construct to be unobservable in principle, as some writers do, would drain the term of all meaning; for what construct is truly unobservable in principle—i.e., could not be observed (if it exists and one knows where to look) with a sufficiently sensitive instrument? Anger, for example, must—if it exists—consist of physical energy or matter of some sort, and hence be observable in principle. Likewise all feelings, sensations, memories, sets, and thoughts.

A widely held view is that constructs are not hypothetical if defined in physical or physiological terms, like Hebb's cell assembly and Kohler's cortical field. In fact, however, some of the crudest and most naive constructs in psychology have been physical or physiological. Thales posited a soul of water, causing all motion, and present even in lodestone. Other Greeks posited souls of air, motes in the air, blood, round atoms, fiery vapor, and fire or heat. Epicurus explained sleep by either the concentration of parts of the soul, or their escape via the pores; images were aggregates of particles continuously emanating from perceived objects. For Zeno sensation was breath passing from the soul to the senses. Hobbes saw pleasure, desire, and love as motions from the head to the heart. Descartes ascribed recall, attention, imagination, wonder, love, hatred, desire, joy and sadness all to the wiggling of the pineal gland and the resulting change in the flow of the animal spirits.

Were these crude constructs any less hypothetical, or any more respectable for being physical or physiological and observable in principle? Does hunger cease to be a construct if we speculate that it consists of a certain hitherto unobserved type of activity of the hypothalamus? A construct supposedly located in a specified part of the body is as hypothetical as one supposedly located "somewhere in the body" if neither has been observed.

The great virtue of hypothetical constructs is their convenience. Since they pervade our everyday speech, they are hard to avoid using inadvertently; and they do save words—but at the expense of accuracy of communication. Since they necessarily entail Type 2 conjectures, they are subject to all the objections to such conjectures. They are also defended by the same arguments as those already examined—and found invalid—in respect to Type 2 conjectures: "indirect confirmability," alleged aid to research, alleged superior explanatory power, and the fact that they abound in physics (to the dismay of many physicists).

A few special advantages are also occasionally claimed for them. For example, Pratt (1945, p. 268) finds it "apparently" useful to the clinician to conceive of inferred forces (such as repression) as having real existence; but he doesn't say why. Rozeboom (1956, p. 257) claims that the excess, undefined meaning of hypothetical constructs may give them "implications or known empirical relations" beyond those of the variables from which they are inferred; but he does not

show how we can ever *know* the implications and empirical relations of an *unknown* meaning.

An original and thoroughgoing defense of hypothetical constructs is presented by Woodworth (1958). He builds his "behavior dynamics" on such constructs, and devotes many pages to justifying their use. Because of his unequalled experience in psychology, and the effort and skill with which he makes his case, his chief arguments are worth considering in detail.

WOODWORTH'S DEFENSE OF HYPOTHETICAL CONSTRUCTS

Woodworth says present behavior is *directly* determined by present stimuli and the present state (largely unobservable) of the organism—not by the past events that produced its present state. In lieu of defining the present state by the events that produced it, he proposes finding a second way to produce the same overt behavior: e.g., control the amount of food eaten by varying time since last feeding and by giving appetizers (p. 4). By controlling all other factors, we might find that X hours of food deprivation plus Y grams of appetizer produce the same food intake as Z hours of food deprivation—a possibly useful empirical finding. But why drag in the construct of hunger? Woodworth does not show how it would profit us in any way to translate his finding into an assertion of equal states of hunger.

He next cites a long series of examples of the usefulness of introspective reports—but without distinguishing the undoubted usefulness of the reports from the doubtful usefulness of inferring that the reported states (constructs) exist. Consider the construct "understanding." A S may be more likely to follow E's instructions if he says he understands them than if he doesn't. But this may be due to his better understanding, greater motivation (from having committed himself), rehearsal of the instructions (due to E's question), or something else. His later success in following the instructions gives no clue as to why he succeeded. To lay his success to his superior understanding, or the failure of other Ss to lack of understanding, would be unwarranted, unnecessary, and—which is crucial—of no value to E. What matters to E is the correlation between Ss' answers and their later performance. It is this correlation that gives Ss' answers their predictive value—not the inferences E makes from them regarding the presence or absence of the construct "understanding."

Similarly, the usefulness of Ss' reports in psychophysical experiments (which Woodworth makes much of) is not increased if E infers that S perceives what he reports. The results are equally satisfactory if Ss' reports are treated simply as verbal behavior that is correlated with stimulus variations.

If S reports that he acted because of thirst, Woodworth sees no reason to reject his report as untrustworthy (p. 9). But no one proposes rejecting his report *as an empirical datum*. The issue is, why treat

the report as more than that? Why infer at all—either that S is or isn't thirsty? Woodworth has no answer. He notes that Ss who cannot recall their last drink can always report their present state of thirst. This argues for asking S how thirsty he is, rather than when he last drank—but not for inferring thirst in Ss who say they are thirsty.

Woodworth notes that: (1) reported thirst correlates with a physiological state; (2) heat or shock above the pain threshold is fairly uniformly reported as painful; and (3) aches and pains reported to doctors and dentists are usually helpful in locating whatever is wrong in the body. This shows the validity of reports as indicants of *certain observable correlates*—but not the usefulness of inferring unobserved constructs from such reports.

Woodworth would not infer pain in shocked or injured subhuman animals from pain-suggestive motor behavior (p. 10); yet he infers human pain from verbal behavior. He argues the reality of pain from the alleged fact of its relief or stoppage by physical means. But we do not know that actual pain is *ever* relieved or stopped—or ever starts! We observe only *reports* of pain and its relief, and their correlation with overt events. Here again—as in his treatment of the correlation between behavior and reported memories and expectations (p. 13)—Woodworth gives no evidence of the validity or usefulness of inferring hypothetical constructs.

He notes that people report perceiving an environment and doing things to it, rather than receptor excitation and effector functioning, and that such reports usually match the observed facts. From this he infers an unobserved "translation" process between the sensations and the reports, whereby sensory cues are decoded into environmental terms and overt acts are precoded into muscle movements (p. 15). However plausible this inference may be, he fails to show *what psychology gains* by making it. The same comment applies to his inference of a recognition process (p. 31) and his approval of Kohler's inference of an unobserved perceptual organization process (p. 26).

Woodworth says his construct of "translation" is needed to show "how behavior can possibly deal with the environment" (p. 143). Trying to show—*without empirical evidence*—how behavior can effectively deal with the environment would no doubt require *some* hypothetical construct (not necessarily Woodworth's). But why should a scientist *try* to show things without empirical evidence? Hypothetical constructs are admittedly useful in fabricating hypothetical explanations; but psychology would be better off without such explanations.

Woodworth says we "must" bridge the "fatal gap" between the sensory stimulus and the muscular response (p. 142). Why is the gap "fatal," and why bridge it with a construct? To be sure, we should seek new *data* to try to close the gap, since closing gaps in our knowledge is the purpose of scientific research. But why try to close a gap in our knowledge with our imagination? Bridging gaps in this way

may be, to some psychologists, as satisfying as a chess game; but it does not advance the progress of science.

He says we *know* "brain readiness" (set) exists because we can observe muscle readiness and we know muscles are controlled by nerves (p. 41). But motor-nerve readiness is not the same thing as preparatory set to do such things as carry out instructions. He says consideration of preparatory set is indispensable in describing, predicting, and controlling human behavior. Admittedly it is indispensable to consider the *data from which he infers preparatory set*—i.e., antecedent conditions and introspective reports. But inferring preparatory set from these data is not indispensable—or even useful—since S's behavior can be correlated directly (and more easily) with the data.

The core of Woodworth's psychology is the construct "drive"—an inferred internal force disposing the organism to respond (p. 58ff). Since drive can be inferred only from actual responses and reports of desires, purposes, etc., inferring drive is no more useful than just noting the organism's observed response probabilities and introspective reports. The primary correlations are between these data, not the construct, and other observed events. He could have built his psychology on these primary correlations.

Woodworth quotes Skinner's argument (p. 7) that since we have to go back to the data anyway (in order to measure the construct), why not save time—and avoid a dangerous pitfall—by examining behavior as a direct function of the data used to infer the construct? He never answers this question. Although he repeatedly asserts the usefulness of hypothetical constructs, he offers no evidence of their usefulness. The conclusion seems warranted that he adduced no evidence because history contains none that hypothetical constructs are, on balance, useful in psychology.

THE IMPLICATIONS FOR PSYCHOLOGY

If there can never be empirical evidence in support of Type 2 conjectures, and if they divert scientists from more profitable research and add nothing to our knowledge, what are the implications for psychologists? How should we modify our practices?

In diagnosis, prognosis, and therapy, putting any reliance on Type 2 conjectures about particular patients is clearly unjustified. Sarbin (1944, p. 227) condemns the use by trained clinical workers of "unverified" hypotheses—and all Type 2 hypotheses are unverified. Testers and diagnosticians, says Anastasi (1950), should not talk or think in terms of will-o'-the-wisp psychological processes which are distinct from performance (p. 77), nor claim that any test measures anything except its criteria (p. 67). Meehl (1954, p. vii) suspects that trying to do selective and prognostic jobs by dynamic formulations and staff conferences wastes much clinical time, since such jobs can often be better done statistically—i.e., without Type 2 conjectures.

Meehl sees merit in the common clinical practice of using Type 2 conjectures as working hypotheses to suggest lines of questioning to the interviewer or therapist. While Type 2 conjectures, so used, may be better than none, Type 1 conjectures, having at least a trace of empirical validity, should give better results.

Research psychologists who are sure that Type 2 concepts help them in their thinking would be as loath to give them up as a person learning to swim would be loath to give up the water wings he is sure are helping him. In both cases the felt usefulness may be illusory. Still, for many of us the habit of Type 2 thinking may be virtually unbreakable. What, then, can we do?

We might begin by resolving never to refer to experiments as tests of Type 2 conjectures, since only Type 1 conjectures can be confirmed, and never to cite the results of published experiments to support or refute Type 2 conjectures. Science is not advanced, and much time is wasted, by arguing the truth-value of Type 2 conjectures. Discussions and debates of Type 1 conjectures, however, are healthy. Such conjectures offer less latitude than Type 2 for variation in the interpretation of evidence, less room for extended debate. They are apt to be sooner abandoned in the face of adverse evidence. Above all, controversy over them is likely to consist chiefly of presenting and evaluating conflicting experimental results—which is highly desirable. Debates on Type 2 theories, on the other hand, are less easily resolved and more likely to produce voluminous (often emotional) verbiage centered around analogical, anecdotal, introspective, and philosophical "evidence."

We could certainly stop organizing textbooks, courses, theses, and papers around Type 2 concepts. To develop high-order laws in psychology, comparable to Newton's laws of motion, we need to scrutinize the entire range of reported overt behavioral data, and try to detect and abstract common properties in widely differing forms of behavior. This abstraction process will be aided if we start now to organize our thinking and writing, and classify our data, on the basis of overt characteristics of behavior. Davis (1953), Bolles (1957), and Turner (1961) have already proposed the reconstruction of psychology along these lines.

Finally, if we can't resist publishing Type 2 conjectures, we can avoid implying their probability or testability. Well-established, allegedly testable Type 2 theories provide ready-made, fashionable molds into which experimentation can be poured, thus inviting channelization of research and diversion of experimenters from systematic investigation of empirical relations. These dangers can be minimized by offering Type 2 theories merely as "queries" (a la Newton), "speculations," or "possibilities."

SUMMARY

Inestimable confusion has resulted from the common practice of

discussing the role of theory in science without specifying the kind of theory referred to: i.e., conjectures based on abstraction from past sensory observations, which are the backbone of science, or Type 2 conjectures, which postulate imagined properties. The latter can never receive empirical support (by "indirect confirmation" or otherwise); they impede the progress of science; they contribute nothing to scientific knowledge; and they cause much waste of time, money, and journal space in futile debates. Hypothetical constructs—giving out "a dream of our own imagination for a pattern of the world"—serve no useful purpose in science.

REFERENCES

ADAMS, D. K. Note on method. *Psychol. Rev.* 1937, 44, 212-218.
ANASTASI, ANNE. The concept of validity in the interpretation of test scores. *Educ. psychol. Measmt,* 1950, 10, 67-78.
APA Committee on Psychological Tests. Technical recommendations for psychological tests and diagnostic techniques. *Psychol. Bull. Suppl.* 1954, 51, 2, Part 2, 1-38.
BACON, F. *The great instauration: the plan of the work.* 1620. Reprinted in *Selected writings of Francis Bacon; with an introduction and notes by Hugh G. Dick.* New York: Modern Library, 1955. Pp. 439-451. (a)
BACON, F. *The new organon.* 1620. Reprinted as cited above. Pp. 453-540. (b)
BECHTOLDT, H. B. Construct validity: a critique. *Amer. Psychologist,* 1960, 15, 619-629.
BERGMANN, G. *Philosophy of science.* Madison: Univer. Wisconsin Press, 1957.
BERGMANN, G., and SPENCE, K. W. Operationism and theory in psychology. *Psychol. Rev.,* 1941, 48, 1-14.
BEVERIDGE, W. I. B. *The art of scientific investigation.* (Rev. ed.) New York: Norton, 1957.
BLAKE, R. M., DUCASSE, C. J. and MADDEN, E. H. *Theories of scientific method: the Renaissance through the nineteenth century.* Seattle: Univer. Washington Press, 1960.
BOLLES, R. C. Occam's razor and the science of behavior. *Psychol. Rep.,* 1957, 3, 321-324.
BORING, E. G. The role of theory in experimental psychology. *Amer. J. Psychol.,* 1953, 66, 169-184.
BRAITHWAITE, R. B. *Scientific explanation: a study of the function of theory, probability, and law in science.* Cambridge: Cambridge Univer. Press, 1953.
BRIDGMAN, P. W. *The logic of modern physics.* New York: Macmillan, 1927.
BROWN, G. B. *Science: its method and its philosophy.* London: Allen & Unwin, 1950.
CAJORI, F. *Newton's Principia: Motte's translation revised.* Berkeley: Univer. California Press, 1934.
CAMPBELL, D. T., & FISKE, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.,* 1959, 56, 81-105.
CONANT, J. B. *Science and common sense.* New Haven: Yale Univer. Press, 1951.
CRONBACH, L. J., & MEEHL, P. E. Construct validity in psychological tests. *Psychol. Bull.,* 1955, 52, 281-302.
DALLENBACH, K. M. The place of theory in science. *Psychol. Rev.,* 1953, 60, 33-39.
DAVIS, R. C. Physical psychology. *Psychol. Rev.,* 1953, 60, 7-14.
DINGLE, H. *Science and human experience.* London: Williams & Norgate, 1931.

DUHEM, P. *La théorie physique: son objet et sa structure.* Paris: Marcel Rivière, 1914.

ENGLISH, H. B., & ENGLISH, A. C. *A comprehensive dictionary of psychological and psychoanalytical terms: a guide to usage.* New York: Longmans, Green, 1958.

FARRELL, B. A. On the limits of experimental psychology. *Brit. J. Psychol.,* 1955, 46, 165-177.

FEIGL, H. Philosophical embarrassments of psychology. *Amer. Psychologist,* 1959, 14, 115-128.

GAMOW, G. *Matter, earth, and sky.* Englewood Cliffs, N. J.: Prentice-Hall, 1958.

GELDARD, F. A. 'Explanatory principles' in psychology. *Psychol. Rev.,* 1939, 46, 411-424.

GUTHRIE, E. R. On the nature of psychological explanations. *Psychol. Rev.,* 1933, 40, 124-137.

HARLOW, H. F. Motivation as a factor in the acquisition of new responses. In M. R. Jones (Ed.), *Current theory and research in motivation: a symposium.* Vol. 1. Lincoln: Univer. Nebraska Press, 1953. Pp. 24-49.

HEMPEL, C. G. *Fundamentals of concept formation in empirical science.* Chicago: Univer. Chicago Press, 1952.

HERSCHEL, J. F. W. *Preliminary discourse on the study of natural philosophy.* (Cabinet cyclopaedia ed.) London: Longman, Brown, Green and Longmans, 1842.

HULL, C. L. The conflicting psychologies of learning—a way out. *Psychol. Rev.,* 1935, 42, 491-516.

HULL, C. L. Mind, mechanism, and adaptive behavior. *Psychol. Rev.,* 1937, 44, 1-32.

JENKINS, J. J., & LYKKEN, D. T. Individual differences. *Ann. Rev. Psychol.,* 1957, 8, 79-112.

JEVONS, W. S. *The principles of science: a treatise on logic and scientific method.* New York: Macmillan, 1874.

JOHNSON, H. M. On verifying hypotheses by verifying their implicates. *Amer. J. Psychol.,* 1954, 67, 723-727.

KENDLER, H. H. "What is learned?"—a theoretical blind alley. *Psychol. Rev.* 1952, 59, 269-277.

KESSEN, W., & KIMBLE, G. A. "Dynamic systems" and theory construction. *Psychol. Rev.,* 1952, 59, 263-267.

KNEALE, W. *Probability and induction.* 1949. Pages 92-110 reprinted in H. Feigl & May Brodbeck (Eds.), *Readings in the philosophy of science.* New York: Appleton-Century-Crofts, 1953. Pp. 353-367.

KOCH, S. Behavior as "intrinsically" regulated; work notes toward a pre-theory of phenomena called "motivational." In M. R. Jones (Ed.), *Current theory and research in motivation: a symposium.* Vol. 4. Lincoln: Univer. Nebraska Press, 1956. Pp. 42-87.

KRECH, D. Dynamic systems, psychological fields, and hypothetical constructs. *Psychol. Rev.,* 1950, 57, 283-290.

LEWIN, K. Behavior and development as a function of the total situation. In L. Carmichael, *Manual of child psychology.* (2nd ed.) New York: Wiley, 1954. Pp. 918-970.

LOEVINGER, JANE. Objective tests as instruments of psychological theory. *Psychol. Rep.,* 1957, 3, 635-694. (*Monogr. Suppl.* No. 9).

MacCORQUODALE, K., & MEEHL, P. E. On a distinction between hypothetical constructs and intervening variables. *Psychol. Rev.,* 1948, 55, 95-107.

McGUIGAN, F. J. Confirmation of theories in psychology. *Psychol. Rev.,* 1956, 63, 98-104.

MARX, M. H. Intervening variable or hypothetical construct? *Psychol, Rev.,* 1951, 58, 235-247.

MAZE, J. R. Do intervening variables intervene? *Psychol. Rev.,* 1954, 61, 226-234.

MEEHL, P. E. *Clinical versus statistical prediction.* Minneapolis: Univer. Minnesota Press, 1954.

MILL, J. S. *System of logic.* (8th ed.) New York: Harper, 1874.

MILLER, J. G. Symbolic technique in psychological theory. *Psychol. Rev.*, 1939, 46, 464-479.

O'NEIL, W. M. Hypothetical terms and relations in psychological theorizing. *Brit. J. Psychol.*, 1953, 44, 211-220.

PIERCE, C. S. Illustrations of the logic of science. Sixth paper. Deduction, induction, and hypothesis. *Popular Sci. Mon.*, 1878, 13, 470-482.

POLYA, G. *Mathematics and plausible reasoning.* 2 vols. Princeton: Princeton Univer. Press, 1954.

PRATT, C. C. Operationism in psychology. *Psychol. Rev.*, 1945, 52, 262-269.

RANKINE, W. V. M. Outlines of the science energetics. *Proc. Philosophical Soc. Glasgow,* 1855, vol. 3. Reprinted in W. V. M. Rankine, *Miscellaneous scientific papers.* London: Griffin, 1881. Pp. 209-228.

REICHENBACH, H. *The theory of probability: an inquiry into the logical and mathematical foundations of the calculus of probability.* (2nd ed.) Trans. by E. Hutten & Maria Reichenbach. Berkeley: Univer. California Press, 1949.

RITCHIE, B. F. A logical and experimental analysis of the laws of motivation. In M. R. Jones (Ed.), *Current theory and research in motivation: a symposium.* Vol. 2. Lincoln: Univer. Nebraska Press, 1954. Pp. 121-176.

ROZEBOOM, W. W. Mediational variables in scientific theory. *Psychol. Rev.*, 1956, 63, 249-264.

RUSSELL, B. *Mysticism and logic and other essays.* New York: Longmans, Green, 1918.

RUSSELL, B. *Analysis of matter.* New York: Harcourt, Brace, 1927.

RUSSELL, B. *An inquiry into meaning and truth.* New York: Norton, 1940.

SARBIN, T. R. The logic of prediction in psychology. *Psychol. Rev.*, 1944, 51, 210-228.

SARBIN, T. R. Role theory. In G. Lindzey (Ed.), *Handbook of social psychology.* Vol. 1. *Theory and method.* Reading, Mass.: Addison-Wesley, 1954. Pp. 223-258.

SCHLICK, M. *Philosophy of nature.* Trans. by A. von Zeppelin. New York: Philosophical Library, 1949.

SIDMAN, M. *Tactics of scientific research: evaluating experimental data in psychology.* New York: Basic Books, 1960.

SKINNER, B. F. Are theories of learning necessary? *Psychol. Rev.*, 1950, 57, 193-216.

SKINNER, B. F. *Science and human behavior.* New York: Macmillan, 1953.

SKINNER, B. F. *Verbal behavior.* New York: Appleton-Century-Crofts, 1957.

STEVENS, S. S. Psychology and the science of science. *Psychol. Bull.*, 1939, 36, 221-263.

TOLMAN, E. C. *Purposive behavior in animals and men.* New York: Century, 1932.

TORGERSON, W. S. *Theory and methods of scaling.* New York: Wiley, 1958.

TURNER, W. S. Can psychology be made wholly objective? *Amer. Psychologist,* 1961, 16, 142.

UNDERWOOD, B. J. *Psychological research.* New York: Appleton-Century-Crofts, 1957.

WHEWELL, W. *Novum organum renovatum.* London: John W. Parker & son, 1858.

WHITTAKER, E. Eddington's principle in the philosophy of science. *Amer. Scientist,* 1952, 40, 45-60.

WOODGER, J. H. *Physics, psychology and medicine. A methodological essay.* Cambridge: Cambridge Univer. Press, 1956.

WOODWORTH, R. S. *Dynamics of behavior.* New York: Holt, 1958.

ZUBIN, J. On the powers of models. *J. Pers.*, 1952, 20, 430-439.

MOUSE EXPLORATORY BEHAVIOR
AND BODY WEIGHT[1]

D. D. THIESSEN
University of California

Thompson (1953) and McClearn (1959) have reported on genetic factors in mouse exploratory behavior. Thompson systematically tested 14 inbred strains and 1 hybrid strain of *Mus musculus* for food drive, emotionality, and exploratory behavior. Large interstrain differences were found on all tests; food drive and emotionality were negatively correlated and no relationship was found between food drive and exploratory activity. McClearn was successful in replicating Thompson's results and demonstrated in a variety of situations that strain differences in activity in novel situations are not specific to any one technique of measurement but suggest a "fundamental mouse personality."

Although Thompson and McClearn were more interested in exploratory behavior than activity as such, Richter's (1958) finding that obesity in rats is a characteristic associated with low activity may have some important implications for their measures of exploration. The hypothesis tested here is that body weight will be inversely related to the extent of exploration for those animals or strains tending toward obesity. Also to be studied will be the contribution of emotionality to strain differences in exploratory behavior and the relevance of emotionality to any behavior changes that may take place over time.

METHOD

Subjects

Three strains of mice were chosen for study: C57BL, BALB/c, and C3H/CRGL/2.[2] Thirty males of each strain were obtained from Cancer Research Genetics Laboratory, University of California, Berkeley Campus at approximately 60 days of age. They were approximately 130 days old when tested in this experiment. About 60 days prior to this time all animals had been tested for learning on a Y maze under conditions of hunger. Subgroups within each strain learned under three different stimulus conditions, i. e., visual, tactual, and auditory.

[2] The C57BL and C3H/CRGL/2 strains are genetically similar to the C57 and C3H strains referred to later in this paper. Origin, degree of inbreeding, and tumor characteristics of these strains have been described (Committee on Standardized Nomenclature, 1960). The C3H/CRGL/2 strain will hereafter be referred to as C3H/2.

Colony Conditions

Three animals were housed per cage, one from each of the three strains represented. This condition existed since weaning. Diablo food and water were available in the living cages at all times. An artificial light cycle was maintained, with lights on from 7 A.M. to 7 P.M.

Apparatus

The apparatus used is a modification of the arena designed by McClearn (1959). It consists of a floor 29⅛ by 29⅛ in. subdivided into 36 squares measuring 4⅞ by 4⅞ in. Barriers 4⅞ in. sq. and ⅛ in. thick are erected at staggered intervals on the square borders of the floor. The floor is enclosed by walls 35 in. high at the back sloping to 25 in. at the front. Doors opening at the front complete the enclosure. Illumination is provided from four 10-w, 115-v frosted bulbs, one in each corner 23 in. from the floor. A transparent plastic box hinged at barrier height in the front left square serves as a starting box. It is controlled by a string leading through the top of the box. The sloped top contains a one-way mirror. The floor and barriers are painted olive drab and other parts light grey.

Procedure

The mouse was weighed to the nearest decigram, placed in the starting box, doors closed, starting box raised, and a timer started. The number of square entries was recorded on a hand tally counter during a 3 min. period. The count at each successive minute was noted on a record sheet. A "square entry" was defined as the placing of both fore-paws across a border. At the end of the 3 min. subject was retrieved, presence of urine noted, and number of fecal boli counted. The boli were removed and the floor wiped clean with a moist sponge.

Testing Schedule

One trial was given to each subject, testing being conducted be-

TABLE 1

SQUARE ENTRY and BODY WEIGHT RESULTS, ALL STRAINS

(N=30 per Strain)

Measure	Strain Mean			Analysis of Variance			t test (df=29)		
	C57BL	C3H/2	BALB/c	Mean Square			C57BL vs. C3H/2	C57BL vs. BALB/c	C3H/2 vs. BALB/c
				Betwn (df=2)	Within (df=87)	F			
Sq. Entry	129.5	94.1	91.9	13342.00	1523.00	8.76**	3.40**	4.53**	0.20
Body Wt.	28.9	33.5	31.0	157.35	13.04	12.07**	3.77**	1.88	2.81*

* p<.05   ** p<.01

tween 1 P.M. and 3 P.M. in a windowless room. The sequence of running was determined from a table of random numbers.

RESULTS

Table 1 summarizes mean, overall F, and t values for body weights and square entries during the 3-min. exploratory period.

Table 2 indicates the number of mice defecating and urinating during the 3-min. testing period. Chi square was computed for independent samples. Product-moment correlations between body weight and 3-min. arena score are also given in Table 2.

Nonparametric statistics were computed according to Siegel (1956) and parametric statistics according to Edwards (1950). An N of 30 for each mouse strain is indicated in the tables.

TABLE 2

DEFECATION, URINATION, and CORRELATIONAL RESULTS, ALL STRAINS

(N=30 per Strain)

Strain	No. Defecating	Chi Square (df=2)	No. Urinating	Chi Square (df=2)	r, Body Weight & Arena Score
C57BL	11		3		.01
C3H/2	10	16.28***	0	11.78**	-.38*
BALB/c	24		9		.06

*p<.05   **p<.01   ***P>.001

These measures show: (a) C57BL strain makes significantly more square entries during the 3-min. period than either the C3H/2 or BALB/c strains, the C3H/2 and BALB/c strains not differing between themselves; (b) body weight for the C3H/2 strain is significantly greater than the other two strains, which do not differ between themselves; (c) defecation and urination scores show that significantly more BALB/c animals defecate and urinate during the 3-min. testing period; and (d) of the three strains the C3H/2 shows a significant inverse correlation between body weight and arena score.

McNemar Test for Significance of Change for related samples was made between successive minutes of the 3-min. exploratory period. Number of animals exhibiting a change in activity between successive min. are given in Table 3. The C57BL strain shows a significant change in activity from minute to minute. More of these animals have

higher scores during the second minute than during the first. No significant changes occur between minutes 2 and 3 for any strain.

TABLE 3

ACTIVITY CHANGE OVER SUCCESSIVE MINUTES, ALL STRAINS

(N=30 per Strain)

Strain	Minute 2			Minute 3		
	N Higher Than on Min. 1	N Lower Than on Min. 1	Chi Square (df=1)	N Higher Than on Min. 2	N Lower Than on Min. 2	Chi Square (df=1)
C57BL	22	7	6.76*	12	18	1.17
C3H/2	20	9	3.45	14	14	0.00
BALB/c	15	15	0.00	15	12	0.14

*$p < .05$

DISCUSSION

The chief findings are that the three strains studied differ in the amount of arena activity, in body weight, in the relationship of arena activity to body weight, in defecation and urination scores, and in the variation of activity from minute to minute. The genetic stability of these differences is indicated by the fact that animals from all three strains lived together under equated environmental conditions since weaning. How one mouse strain scores on these variables appears to be independent of how any other strain will score, i.e., knowing the pattern of performance on all measures for one strain will not help in the prediction of the pattern of performance for any other strain. In this sense there is not only a "fundamental mouse personality" as McClearn suggests, characterizing each strain across a variety of behavior situations, but in addition there is a unique pattern of performances for each strain in a single environmental situation.

Note that activity scores and body weights are inversely correlated for the C3H/2 mouse strain. Lyon, Dowling, and Penton (1953a, b) found that many animals of the C3H strain, of the same age as those used in this study, became obese without special diet or manipulation. The C57 strain, however, did not show this tendency (the BALB/c strain used in this study was not tested by Lyon, et al.). Lyon, et al. attributed these findings to the innate differences between the strains in the ability to oxidize fat and fat-producing substances. On a high fat diet the C3H strain consumed less oxygen and deposited more carcass fat than did the C57 strain. These findings coupled with Richter's (1958)

finding that obesity is related to lowered activity suggest that the inverse relationship found between body weight and exploratory behavior for the C3H/2 animals in this study resulted from the obeseness that many of these animals displayed. The C57BL strain showed neither obeseness nor the inverse correlation between body weight and exploratory behavior. In these respects the BALB/c strain was more comparable to the C57BL strain than to the C3H/2 strain. These observations indicate that morphological and physiological variables are an important consideration in the analysis of exploratory behavior, particularly when using a comparative approach where it cannot be assumed that all developmental sequences, behavioral or biological, proceed at similar rates or are even functionally equivalent.

Change in response to the novel situation appears to be taking place over the three-min. period for at least one strain. Significantly more animals of the C57BL strain scored higher in activity during the second minute than during the first. Twenty out of 30 C3H/2 animals also scored higher during the second minute, although this change is not statistically significant. The BALB/c strain was virtually constant in its activity over the three minutes measured. Note that the BALB/c strain was both highest in emotionality (defecation and urination scores) and lowest in the number of square entries. Conversely, the C57BL and C3H/2 strains were lower in emotionality and higher in square entries. All of these relationships were statistically significant. It therefore appears that emotionality may be the common denominator governing exploratory behavior among these strains; greater exploratory behavior and change in response to novelty appear to be contingent upon low emotionality, whereas less exploratory behavior and little or no behavior change is contingent upon high emotionality. This is consistent with the observations of Montgomery and Monkman (1955) that induced "fear" can markedly reduce exploratory behavior.

These results point out the importance of strain differences, biological development, and emotionality as possible factors influencing exploratory behavior.

SUMMARY

Three inbred strains of mice were tested in a modified open field apparatus. Even though these strains were housed together since weaning they were found to differ in exploratory activity, body weight, relationship to activity to body weight, and defecation and urination scores. Body weight was seen to be inversely related to activity for one strain. Obesity of these animals was suggested as being the instrumental factor accounting for this relationship. A strain of animals showing high activity and low emotionality also showed an increase in activity over the three-min. test period. Another strain showing low activity and high emotionality did not show this adaptation over successive minutes. The third strain appeared to be somewhat intermediate on these measures.

From this the suggestion was drawn that the primary factor common to all three strains governing the amount of exploration is the extent of the emotional response.

REFERENCES

Committee on Standardized Nomenclature for Inbred Strains of Mice. Standardized nomenclature for inbred strains of mice. *Cancer Res.*, 1960, 20, 145-169.

EDWARDS, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.

LYON, J. B., DOWLING, M. T., & PENTON, P. F. Studies on obesity: I. Nutritional obesity in mice. *J. Nutrition,* 1953, 49, 319-331. (a)

LYON, J. B., DOWLING, M. T., & PENTON, P. F. Studies on obesity: II. Food intake and oxygen consumption. *J. Nutrition,* 1953, 51, 65-71. (b)

McCLEARN, G. E. The genetics of mouse behavior in novel situations. *J. comp. physiol. Psychol.,* 1959, 52, 62-67.

MONTGOMERY, K. C., & MONKMAN, J. A. The relation between fear and exploratory behavior. *J. comp. physiol. Psychol.,* 1955, 7, 145-155.

RICHTER, C. P. Neurological basis of responses to stress. In G. E. W. Wolstenholme (Ed.), *Neurological basis of behavior.* Boston: Little Brown, 1958. Pp. 204-222.

SIEGEL, S. *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill, 1956.

THOMPSON, W. R. The inheritance of behaviour: behavioural differences in fifteen mouse strains. *Canad. J. Psychol.,* 1953, 7, 145-155.

FIELD - ARTICULATION IN RECALL[1]

RILEY W. GARDNER and ROBERT I. LONG
The Menninger Foundation

In a recent publication, Gardner, Holzman, Klein, Linton, and Spence (1959) described a Field-Articulation factor including scores for such tests of "Field-Dependence" as Witkin's Embedded Figures and Rod and Frame Tests (Witkin, Lewis, Hertzman, Machover, Meissner, and Wapner, 1954). Gardner et al. suggested that the commonality linking responses to these superficially different tests was the requirement that S attend selectively to relevant vs. irrelevant cues. The usefulness of this interpretation of the response processes involved in these tests has achieved support from a subsequent study (Gardner, Jackson, and Messick, in press), in which successful predictions were made from positions along the Field-Articulation dimension to performances in tests of several intellectual abilities. The findings of Gardner, Jackson and Messick included relationships between Field-Articulation measures and performance in tests of the intellectual ability long called "Associative Memory." These results led to the present study, which was also suggested by Gollin and Baron's (1954) finding of significant relationships between the speed with which Ss identified simple figures in an embedded figures test and amount recalled and rate of relearning in a retroactive inhibition experiment involving nonsense syllables. These investigators found no relationship between this score for an embedded figures test and learning measures.

Gollin and Baron offered several tentative hypotheses concerning the relationship between speed of identifying embedded figures and amount recalled or rate of relearning in a retroactive inhibition situation. At two points, they suggested that these individual consistencies represent the (p. 259) "facility with which a figure can be maintained against a ground," or (p. 261) the "stability with which previously experienced material is maintained under conditions of interference." At other points, they suggested that these consistencies represent (p. 261) "the facility with which figural material could be produced in the face of hindering field conditions," and that "motivational variables may contribute to figure-ground facility." Their final comment (p. 261) concerning "ability to make distinctions between aspects of given field" is most similar to the conception of Field-Articulation guiding the present study.

The present study includes Witkin's Embedded Figures Test and

Rod and Frame Test to provide measures of selectivity in attention deployment. Predictions are made to individual differences in recall and recognition of words under conditions of interference. It was assumed that the selective deployment of attention required for rapid identification of simple figures in masking contexts is similar to the requirement to attend selectively to gravity cues and the rod (but not the frame) in the Rod and Frame Test, and the requirement to attend selectively to elements of memory organizations representing two lists of similar words. In the last of these tests, recall was presumed to involve selective deployment of attention to relevant vs. irrelevant cues in two ways: in the requirement to recall words in order under conditions of high within-list similarity; and in the simultaneous requirement to identify which of two highly similar lists contained each word recalled.

METHOD

Subjects

The Embedded Figures and Rod and Frame Tests were originally administered as part of a larger battery of procedures to 80 female students at a small midwestern university. Three years later, the 38 Ss still in the community were retested on a variety of procedures and given the Word Recall Test described here. At the time of the original testing, they ranged in age from 18 to 22, mean age, 19.3. During the three-year interval, 19 of the 38 Ss had been graduated, 13 had been married, and 4 had become pregnant.

Procedures

Embedded Figures Test. Witkin's Embedded Figures Test was administered as Witkin (1950) suggests. S is required to identify one of eight simple figures in each of 24 complex figures. The score used was the total time to identify the simple figures in the 24 items.

Rod and Frame Test. This test was administered as suggested by Witkin et al. (1954). S, sitting in a chair in a dark room, views a luminous rod contained in a luminous square frame. On each of the 24 trials, rod and frame are tilted 28° right or left, in the same or opposite directions. S directs E how to adjust the rod to make it truly vertical in three sets of eight trials under these conditions: (a) body tilted left 28°; (b) body tilted right 28°; (c) body erect. The score is average error for the 24 trials.

Word Recall Test. Two lists of similar words were presented. List A was "Printed, Painted, Peppered, Patted, Plated, Pounded, Peddled, Pictured." List B was "Posted, Prodded, Planted, Petted, Powdered, Parted, Pleated, Papered." Both recall and recognition measures were obtained. The recognition list consisted of 25 words (the 16 stimulus words and 9 other words beginning with "P"). The procedure was as follows:

E said, "I am going to read a short list of words to you. Please lis-

ten carefully and concentrate because I want you to remember as many of these words as you can. I shall read the list through three times in the same order."

E then read List A three times at the rate of one word every two seconds, with a six-second pause between readings.

Immediately after the third reading of List A, E said, "Now I am going to read another list three times. Please listen carefully and concentrate, because I also want you to remember as many words as you can in this list."

E read List B three times at the rate used for List A.

After the third reading of List B, E passed out one blank sheet of lined paper numbered from one to eight down the left-hand margin and said, "Please put your name in the upper right-hand corner of this sheet. Now I would like you to write all the words you can remember that were in the first list I read. When in doubt about a word or about its position in the list, write it down anyway. Please do not erase. If you wish to remove a word or change its position draw a line through it."

E then passed out the sheets for List B and repeated the above instructions.

After S finished List B, the recognition sheet was passed out and E said, "Please write your name on this sheet. All the words in the two lists I read are on this sheet. In addition, there are some other words which were not in either of the two lists. Put a '1' in the blank in front of each word that you think was on the first list and a '2' in front of each word that you think was on the second list. If you think the word was on neither of the two lists, leave it blank. I want you to mark a '1' in front of eight words and a '2' in front of eight other words, even if you have to guess at some. Do not erase. If you wish to change an answer draw a line through it. Go ahead."

Recall scores were the number in correct position, correct list, for List A, List B, and A plus B. Recognition scores were the number correct for List A, List B, and A plus B.

RESULTS

For the 38 Ss, the mean numbers in correct position, correct list, on recall were A, 3.0; B, 2.4. The mean numbers correct in the recognition portion of the test were A, 5.8; B, 5.1. The r between the total recall and recognition scores was .43 ($p<.01$). The r between the Embedded Figures and Rod and Frame Test scores was .55 ($p<.01$), for the 26 Ss who took both tests.

TABLE 1
PEARSON CORRELATIONS

	Embedded Figures Test Total Time ($N = 26$)	Rod and Frame Test Average Error ($N = 38$)
Word Recall		
List A	−.43°	−.28
List B	−.31	−.29
Lists A + B	−.54°°	−.40°
Word Recognition		
List A	−.43°	−.17
List B	−.21	−.27
Lists A + B	−.46°	−.27

°$p<.05$.   °°$p<.01$.

All the r's in Table 1 are in the predicted direction, although only one of those involving the Rod and Frame Test score is significant. The r's are somewhat higher for List A than for List B. In the present sample, the r's tend to be somewhat higher for the total recall and recognition scores than for the scores derived from individual lists, possibly because of the limited ranges of scores for the individual lists.

DISCUSSION

Our results with List A confirm the hypothesis that the speed with which persons identify embedded figures is significantly related to the number of items correct in recall under conditions of interference despite administration of the Word Recall Test three years after the Embedded Figures Test. In this connection, it should be pointed out that Witkin et al. (1954, p. 86) report a test-retest r of .89 for the Embedded Figures Test for an interval of approximately three years. The fact that only one of the r's between the Rod and Frame Test score and our recall and recognition measures is significant, although all are in the predicted direction, may be attributable to lower reliability of this score as compared to the Embedded Figures Test score; to the fact that it provides a less adequate measure of individual differences in selectivity of attention (suggested by the somewhat lower loadings for the Rod and Frame Test score on the Field-Articulation factors described by Gardner et al., 1959; in press); or to the absence of a speed factor in this test which could be associated with rate of learning the two lists of words.

The principal hypothesis concerning individual differences in selectivity of attention that led to the present study can be summarized as follows:

A. "Extraction" of embedded simple figures requires selective deployment of attention during scanning of the complex figure. "Percep-

tual recognition" of a number of these figures is clearly impossible for many Ss (Gardner et al., 1959), who must "trace" them visually in order to identify them. The significant correlations (see Gardner et al., 1959) between solution times for easy and difficult figures suggests that selective attention to the relevant parts of the complex figures may be the key to successful response in both easy and difficult items.

B. The Rod and Frame Test requires that S attend selectively to the only available cues that can lead to effective estimation of the verticality of the rod—the gravity cues—and to the rod, while actively withholding attention and response from the misleading, irrelevant frame.

C. The Word Recall Test sets up an original memory organization representing List A and a second memory organization representing List B. When S is asked to recall List A, the memory organization representing List B serves as irrelevant, "masking" material that increases the difficulty of recall or recognition. In order to perform effectively, S must attend selectively to elements of the memory organization representing List A, and not attend to interfering memories from List B. When asked to reproduce List B, he must reverse the pattern of selective scanning of the two memory organizations. When both memory organizations have been formed, response to either alone can be conceived of as an internal "embedded figures" test analogous to the usual embedded figures tests, in which both relevant and irrelevant stimuli are external to S. It is in this sense that these three superficially very different tests—Embedded Figures, Rod and Frame, and Word Recall—may be related.

In the present design, however, the effects of selectivity of attention on learning and recall could not be distinguished. A further study is indicated in which Ss are required to learn each of the two lists to a criterion. If the relationships observed in the present study hold under those conditions, it will be clear that the individual differences in selectivity of attention indicated by scores for Witkin's tests are associated with selectivity in recall and recognition rather than learning per se. Such a study would also allow a test of the reverse hypothesis—that selectivity primarily affects rate of learning in the Word Recall Test.

SUMMARY

Performance on Witkin's Embedded Figures and Rod and Frame Tests were related to recall and recognition of two lists of words in an interference situation. The results support Gollin and Baron's (1954) findings that speed of locating embedded figures is related to recall and recognition under these conditions. The results are discussed in terms of individual differences in the capacity to attend selectively to relevant vs. irrelevant material.

REFERENCES

GARDNER, R. W., HOLZMAN, P. S., KLEIN, G. S., LINTON, HARRIET B., and SPENCE, D. P. Cognitive control: A study of individual consistencies in cognitive behavior. *Psychol. Issues,* 1959, 1, No. 4.

GARDNER, R. W., JACKSON, D. N., and MESSICK, S. J.   Personality organization in cognitive controls and intellectual abilities. *Psychol. Issues,* in press.

GOLLIN, E. S., and BARON, A. Response consistency in perception and retention. *J. exp. Psychol.,* 1954, 47, 259-262.

WITKIN, H. A. Individual differences in ease of perception of embedded figures. *J. Pers.,* 1950, 19, 1-15.

WITKIN, H. A., LEWIS, HELEN B., HERTZMAN, M., MACHOVER, KAREN, MEISSNER, PEARL B., and WAPNER, S. *Personality through perception.* New York: Harper, 1954.

PERSPECTIVES IN PSYCHOLOGY

XVIII.   SOME REFLECTIONS ON PERCEPTION[1]

N. H. PRONKO

University of Wichita

When, in the dramatic bedroom scene, Hamlet describes the visitation of his father's ghost, to his mother, Gertrude, she chides her son and attributes the apparition to "the very coinage of your brain."

What a modern ring Shakespeare's explanation has for Hamlet's misperception, and what an ages-long kinship to theories of hallucination and perception that psychologists hold in 1961! If the reader doubts it, let him look in any current textbook of general psychology.

Despite the reaction formations of contemporary psychology in the direction of rigorous experimentation and quantification, our current theories of perception are remarkably similar to those extant in Shakespeare's time. What if basic theory in astronomy, physics, chemistry, and biology were as backward as that of psychology in general and of perception in particular? It is the latter topic that is the focus of the present perspective. Surely, it seems reasonable that theories of perception should have developed beyond Queen Gertrude's surmise that perceptions (or hallucinations) are brain epiphenomena! At any rate, the other sciences have had a number of revolutions of their fundamental theory in the intervening years since Shakespeare wrote Hamlet about 1601. Psychology has been bypassed in this regard and while either Gertrude or the Avon bard would of necessity be speechless among the physical scientists they would nevertheless feel quite at home at Twentieth Century psychologists' convention discussions on the subject of perception, for perceptions are still today believed to be mostly by-products of the brain as they were in Shakespeare's time.

The purpose of the present comments is simply to clear away the dead wood and to move away from the embarrassing impasse of perceptual theory. Let us face it, psychologists are embarrassed about perception. Some shy away from these discriminatory responses and prefer to deal with attending or other (e.g. movement) responses. It is as if the physicist were embarrassed about magnetism or hydrogen sulfide

and disliked dealing with these data. We in psychology are not too confident in the existence of perceptions.

The writer believes that one reason underlying the psychologist's sense of uneasiness with the facts of perception is what might be labelled a "start-from-scratch" or "all-or-nothing" attitude. With this orientation, the psychologist seems to feel that he must give a *complete and exhaustive* account of perceptual data. Is such an attitude really justifiable? Does the physicist do a similar total job for magnetism, cosmic rays or oxygen? To take the last item, oxygen, is this datum really not an altogether evanescent and subtle phenomenon? After all, can the physicist say much more than that this gas is without odor, smell, or taste and that it has a certain property of supporting combustion and a certain weight? And yet, he attacks this problem as he does his others posititistically and vigorously and makes headway in understanding them all. But suppose one had never in his life observed oxygen, magnetism, gravitation, or electricity. Would any amount of description duplicate the experience of these data or substitute for such an experience? Obviously, one would have to observe such events before their theoretical description could have any significance.

Why then do psychologists seem to set up double standards and make excessive demands on their own statements about perception? After all, are not the facts of perception as ubiquitous and abundant as oxygen? Their pervasiveness and shared nature actually give them an advantage over the less intimate and personal conditions in our surroundings. Through social checks we can enter into and compare our perceptions, however subtle they may be, much more intimately than we can ever share or enter into magnetism *qua* magnetism. A realization of this fact should dispel forever our sense of embarrassment over perceptual data and our unnecessary compulsion to describe them exhaustively. After all, they, as much as oxygen and the rest, are the proper starting point for theoretical construction.

Another stumbling block in the development of a valid theory of perception has a semantic origin. In this writer's view, we are too much bound by our everyday language, which distorts and gives a determining direction to theory development. Take the simple situation expressed in the phrasing of the man in the street, "I see a tree."

This statement formulates (nay, straight-jackets) the working basis of current theories of perception. Immediately, an entity, "I," over here and another, "a tree," over there are brought into juxtaposition. So far only two disparate entities have been related as if they were puppets on a stage, but trouble appears when we try to introduce them to the verb, "see." Convention has shied away from imputing "seeingness" to the tree. For that reason, no one has verbally created entities within the tree to carry the theoretical ball. Instead, tradition has always moved

over to the organism's side of the fence and attributed the "seeingness" to a pseudo-localization *within* the organism. From here on, we have double-talk in terms of an "outside," "the tree," and an "inside," the perception *of* the tree.

Several points need to be made concerning the semantic disorder which we have created even before our theory has gotten off the ground. First of all, there is the easily overlooked differential status of the terms "inside" and "outside." The arbitrary carving up of the space referred to can be avoided. Nevertheless, we must acknowledge that even the bad start such a theory has taken should accord a different status to the terrestrial "outside," (i.e. the locus of the participating tree) than it can to the imputed but never designated "inside" the organism. How can we possibly work at establishing the interior location of "the perception?" After all, lens, retina, optic pathways and visual cortex have already been gone over with a fine tooth comb without ever establishing a position for "the percept." What a pity that the epidermis of the organism permitted the erection of this artificial and arbitrary boundary. If we may speculate in science fiction fashion, a more amorphous separation of organism-in-environment might have prevented the theoretical creation of an "inside" and "outside." At any rate, once these non-correlative, non-analogous "spaces" have been rent asunder, no amount of theoretical ingenuity can join them together again.

However, our troubles are not yet over. Another specter rises. This difficulty comes from our traditional predilection for the organism and the resultant need for creating perception *within it*. This preference for the organism creates a *deus ex machina* situation in which (theoretically, at least), I appear to generate "seeingness" even in advance of my confrontation with "tree." The tree appears only as a cue or a trigger for the "seeing." Surely, this is not a correct statement of the reality that we started with.

But it is not only spatial problems that bedevil us, for our "time" is also out of joint according to the manner in which it has been forced to do duty in the construction of perceptual theory. The conventional temporal framework in conventional theories of perception employs *instant* time. And yet no one ever saw tree or anything else for that matter in the same kind of time as is occupied by an explosion. We acknowledge the fact that perceptions *develop over time,* but we do not treat them that way. The writer would venture to predict that our difficulties with perception will not be surmounted until we adopt a framework similar to the one employed by the embryologist or the historical geologist. Their time marches on and is an integral feature of their theoretical description. Perception theory has hobbled along with a cramping, crippling instantaneous temporal framework.

The deleterious side effects of our ages-old perceptual theory have

by no means been exhausted. Consider the looseness and vagueness of perceptual semantics. The term, "perception," refers to (a) the thing perceived, to (b) the perceiving of the thing, to (c) the perception projected back out from the organism's inside to the place where the thing *really is,* and to (d) the mysterious goings-on within the organism. Consider also the obscuring of the undisputed fact and existence of perceptual data by the traditional "eye-as-a-camera" approach implicit in perceptual theory. How the living eye could ever project an image of the tree (even though upside down) on to a transparent retina appears next to impossible with all the imperfections inherent in the visual system. And how the alleged "picture" could be transmitted intact through the optic pathway is a deeper mystery still. But overlook all the preceding and imagine that somehow the tree is imaged on the occipital lobe. How do we get the organism back into the impersonal chain reaction? With what mind's eye could the "image" be seen and what is really being seen,—the tree out there, the picture in the eye, the one on the visual cortex? That theoretical confusion is rampant is rather obvious by this time.

The alternative. In a sketchy way only, the alternative to the foregoing theoretical debauchery is an avoidance of aged cultural impositions or pristine perceptual data. Instead of starting from such popular formulations as *"I see a tree,"* let us attempt an event orientation in which seeing does not arise from the confrontation of an organism and a tree. In fact, organism and tree are only the anchorage points for an occurrence, situation, interaction or transaction that transpires in a larger space-time framework than the traditional "inside" and "outside" the organism and within a more comprehensive temporal structure than the instant. Proper analysis of relevant variables is aided and abetted and such general approaches as those of Kantor, Skinner and Bentley in their different areas of operation are proof of the pudding.

316

BURTON, A. (Ed.) *Psychotherapy of the psychoses.* New York: Basic Books, 1961. Pp. 386.

DELLIS, N. P. & STONE, H. K. *The training of psychotherapists.* Baton Rouge: Louisiana State University Press, 1960. Pp. 195.

DAWSON, J. G., STONE, H. K., & DELLIS, N. P. *Psychotherapy with schizophrenics.* Baton Rouge: Louisiana State University Press, 1961, Pp. 156.

FREUD, S. *The ego and the id.* (James Strachey, tr.) New York: Norton, 1961. Pp. 88.

INDIANA UNIVERSITY PRESS. *Drug addiction: crime or disease?* Bloomington: Author, 1961. Pp. 173.

ODLUM, DORIS. *Journey through adolescence.* Baltimore: Penguin Books, 1961. Pp. 160.

KIMBLE, G. A. *Hilgard and Marquis' conditioning and learning.* New York: Appleton-Century-Crofts, 1961. Pp. 590.

KULP, J. L. (Ed.) *Geochronology of rock systems.* New York: Annals of The New York Academy of Sciences, Vol. 91, Art. 2. Pp. 159-594.

ENGLISH, O. S. et al. *Direct analysis and schizophrenia.* New York: Grune & Stratton, 1961. Pp. 128.

ZIMNY, G. H. *Method in experimental psychology.* New York: Ronald Press, 1961. Pp. 366.

CAPLAN, G. *Prevention of mental disorders in children.* New York: Basic Books, 1961. Pp. 425.

NEWCOMB, T. M. *The acquaintance process.* New York: Holt, Rinehart & Winston, 1961. Pp. 303.

HALLMAN, R. J. *Psychology of literature.* New York: Philosophical Library, 1961. Pp. 262.

JOHNSON, G. O. & BLAKE, KATHRYN A. *Learning performance of retarded and normal children.* Syracuse: Syracuse University Press, 1960. Pp. 216.

BATTISTA, L. A. *Mental drugs: chemistry's challenge to psychotherapy.* New York: Chilton, 1960. Pp. 155.

KLEIN, MELANIE. *Narrative of a child analysis.* New York: Basic Books, 1961. Pp. 496.

BONNER, H. *Psychology of personality.* New York: Ronald, 1961. Pp. 534.

KALINOWSKY, L. B. & HOCH, P. H. *Somatic treatments in psychiatry.* New York: Grune & Stratton, 1961. Pp. 413.

BENNETT, E. *Personality assessment and diagnosis.* New York: Ronald, 1961. Pp. 287.

SINGER, G. *Morale factors in industrial management.* New York: Exposition Press, 1961. Pp. 155.

WYLIE, RUTH C. *The self concept.* Lincoln: University of Nebraska Press, 1961. Pp. 370.

BERNHARDT, K. S. (Ed.) *Training for research in psychology: the Canadian opinicon conference 1960.* Toronto: University of Toronto Press, 1961. Pp. 127.

FULTON, R. B. *Original Marxism—estranged offspring.* Boston: Christopher Publishing House, 1960. Pp. 167.

LOTH, D. *The erotic in literature.* New York: Julian Messner, 1961. Pp. 256.

MUNN, N. L. *Psychology, the fundamentals of human adjustment.* (4th ed.) Boston: Houghton Mifflin, 1961. Pp. 812.

GOLDMAN, L. *Using tests in counseling.* New York: Appleton-Century-Crofts, 1961. Pp. 434.

KAPLAN, B. *Studying personality cross-culturally.* Evanston, Illinois: Row, Peterson, 1961. Pp. 687.

HALL, D. M. *Dynamics of group action.* Danville, Illinois: Interstate, 1960. Pp. 243.

SHIPLEY, T. (Ed.) *Classics in psychology.* New York: Philosophical Library, 1961. Pp. 1342.

EYSENCK, H. J. *Behaviour therapy and the neuroses.* New York: Pergamon, 1960. Pp. 479.

MUENSTERBERGER, W. & AXELRAD, S. (Eds.) *The psychoanalytic study of society.* Vol. I. New York: International Universities Press, 1960. Pp. 384.

CROW, L. D. & CROW, ALICE (Eds.) *Readings in child and adolescent psychology.* New York: Longmans, Green & Co., 1961. Pp. 592.

FRANK, J. D. *Persuasion and healing.* Baltimore: The Johns Hopkins Press, 1961. Pp. 282.

MYERS, F. W. H. *Human personality and its survival of bodily death.* New York: University Books, 1961. Pp. 416.

HALL, J. F. *Psychology of motivation.* Chicago: Lippincott, 1961. Pp. 382.

DUNHAM, H. W. *Sociological theory and mental disorder.* Detroit: Wayne State University Press, 1959. Pp. 312.

DUNHAM, H. W. & WEINBERG, S. K. *The culture of the state mental hospital.* Detroit: Wayne State University Press, 1960. Pp. 309.

DORSEY, J. M. & SEEGERS, W. H. *Living consciously: the science of self.* Detroit: Wayne State University Press, 1959. Pp. 174.

FRAIBERG, L. *Psychoanalysis and American literary criticism.* Detroit: Wayne State University Press, 1960. Pp. 272.

GROUP FOR THE ADVANCEMENT OF PSYCHIATRY. *Symposium No. 7—Application of psychiatric insights to cross-cultural communication.* New York: Author, 1961.

SCHNEIDERS, A. A. *Personality development and adjustment in adolescence.* New York: Grune & Stratton, 1960. Pp. 473.

PASCAL, G. R. & JENKINS, W. O. *Systematic observation of gross human behavior.* New York: Grune & Stratton, 1961. Pp. 126.

MASSERMAN, J. H. *Principles of dynamic psychiatry.* (2nd ed.) Philadelphia: Saunders, 1961. Pp. 332.

MORGAN, C. T. *Introduction to psychology.* (2nd Ed.) New York: McGraw-Hill, 1961. Pp. 727.

SPOTNITZ, H. *The couch and the circle.* New York: Knopf, 1961. Pp. 274.

ACKERMAN, N. W., BEATMAN, F. L., & SHERMAN, S. N. (Eds.) *Exploring the base for family therapy.* New York: Family Service Association of America, 1961. Pp. 159.

VINCENT, E. L. & MARTIN, P. C. *Human psychological development.* New York: Ronald Press, 1961. Pp. 522.

KELLOGG, W. N. *Porpoises and sonar.* Chicago: University of Chicago Press, 1961, Pp. 177.

MASSERMAN, J. H. (Ed.) *Science and psychoanalysis.* Vol. IV. *Psychoanalysis and social process.* New York: Grune & Stratton, 1961. Pp. 196.

DUIJKER, H. C. J. & FRIJDA, N. H. *National character and national stereotypes.* New York: Humanities Press, 1961. Pp. 238.

LEUBA, C. L. *Man: a general psychology.* New York: Holt, Rinehart, & Winston, 1961. Pp. 676.

ANASTASI, ANNE. *Psychological testing.* New York: Macmillan, 1961. Pp. 657.

FREEHILL, M. F. *Gifted children—their psychology and education.* New York: Macmillan, 1961. Pp. 412.

Journal of Individual Psychology

CONTENTS